

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>7</sup> :</b> <b>C12N 15/12, C07K 14/705, 16/28, G01N 33/68, A01K 67/027</b>		<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 00/66735</b>
			<b>(43) International Publication Date:</b> 9 November 2000 (09.11.00)
<b>(21) International Application Number:</b> PCT/US00/12336			<b>(81) Designated States:</b> AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
<b>(22) International Filing Date:</b> 4 May 2000 (04.05.00)			
<b>(30) Priority Data:</b> 60/132,541 5 May 1999 (05.05.99) US			
<b>(71) Applicant:</b> LEXICON GENETICS INCORPORATED [US/US]; 4000 Research Forest Drive, The Woodlands, TX 77381 (US).			
<b>(72) Inventors:</b> TURNER, Alexander, C., Jr.; 67 Winter Wheat Place, The Woodlands, TX 77381 (US). OLSON, Andrew; 43 North Walden Elms Circle, Spring, TX 77382 (US). ZAMBROWICZ, Brian; 18 Firethorne Place, The Woodlands, TX 77382 (US). FRIEDRICH, Glenn, A.; 2207 Hermann Drive, Houston, TX 77004 (US). SANDS, Arthur, T.; 163 Bristol Bend Circle, The Woodlands, TX 77382 (US).			
<b>(74) Agents:</b> CORUZZI, Laura, A. et al.; Pennie & Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036 (US).			<b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
<b>(54) Title:</b> NOVEL HUMAN ION CHANNEL PROTEINS			
<b>(57) Abstract</b> <p>The present invention provides a novel family of human ion channel proteins (ICPs). The invention additionally provides for agonists, antagonists, antibodies, antisense molecules that are specific for the ICPs, and further provides genetically engineered expression vectors for the ICPs and host cells comprising the same. The invention further provides for processes for identifying/producing molecules that effect ICP activity that comprise the use of the disclosed ICPs or genes encoding the same.</p>			

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## NOVEL HUMAN ION CHANNEL PROTEINS

This application claims priority under 35 U.S.C. §119 (e) to U.S. provisional patent application no. 60/132,541 filed May 5, 1999, which is hereby incorporated by reference in its entirety.

5

### **1. INTRODUCTION**

The present invention relates to the discovery, identification and characterization of novel human polynucleotides that encode proteins sharing structural similarity with ion channel proteins. The invention encompasses the described polynucleotides, host cell expression systems, the encoded proteins, fusion proteins, polypeptides and peptides, antibodies to the encoded proteins and peptides, and genetically engineered animals that lack at least one of the disclosed genes, or over express the disclosed genes, antagonists and agonists of the described proteins, and other compounds that modulate the expression or activity of the proteins encoded by the disclosed genes that can be used for diagnosis, drug screening, clinical trial monitoring, and/or the treatment of physiological or behavioral disorders.

### **2. BACKGROUND OF THE INVENTION**

Ion channel proteins selectively facilitate the translocation of ions either into or out of the cell. Such mechanisms play important roles in maintaining cellular and metabolic homeostasis, neuron function, and signaling, and drug resistance. As such, ion channel proteins are natural targets for the development and study of novel therapeutic agents.

### **3. SUMMARY OF THE INVENTION**

The present invention relates to the discovery, identification and characterization of nucleotides that encode novel ion channel-like proteins ICPs, and the corresponding amino acid sequences of the described novel ion channel-like proteins (ICPs). The ICPs described for the first time herein, are membrane associated proteins that are involved in translocating ions across membranes. The described ICPs share structural similarity with a variety of ion channel proteins, and particularly sodium or calcium channel proteins. Similar sodium or

calcium channel proteins can be involved in the voltage dependent ion permeability of excitable membranes.

The sequences encoding the ICPs were initially identified via chimeric gene trap transcripts generated in human cells. The novel human ICPs described herein encode  
5 proteins of about 398 and about 417 amino acids in length (see SEQ ID NOS: 2, and 4 respectively) with at least six distinct hydrophobic domains each. The size of the described proteins is similar to that observed in, for example, sodium channel beta-subunits (attached to sodium channel alpha subunits by disulfide bonds); however, the described proteins share broad similarity with the larger ion channel alpha subunits. As such, the described proteins  
10 describe a novel family of ion channel proteins.

The invention encompasses the nucleotides presented in the Sequence Listing, host cells expressing such nucleotides, the expression products of such nucleotides, and: (a) nucleotides that encode mammalian homologues of the described ICPs including the specifically described human ICP genes; (b) nucleotides that encode one or more portions  
15 that correspond to functional domains of an ICP, as well as the polypeptide products specified by such nucleotide sequences, including but not limited to the novel regions of any extracellular domain(s) (ECD), one or more transmembrane domain(s) (TM), and the cytoplasmic domain(s) (CD); (c) isolated nucleotides that encode mutants, engineered or naturally occurring, of the described ICPs in which all or a part of at least one of the  
20 domains is deleted or altered, and the polypeptide products specified by such nucleotide sequences, including but not limited to soluble receptors in which all or a portion of a TM is deleted, and nonfunctional receptors in which all or a portion of another domain is deleted; (d) nucleotides that encode fusion proteins containing the coding region from an ICP, or one of its domains (*e.g.*, an extracellular domain) fused to another peptide or polypeptide.

25 The invention also encompasses agonists and antagonists of the ICPs, including small molecules, large molecules, mutated ICPs, or portions thereof, and antibodies, as well as nucleotide sequences that can be used to inhibit the expression of the described ICPs (*e.g.*, antisense and ribozyme molecules, and gene or regulatory sequence replacement constructs) or to enhance the expression of the described ICP genes (*e.g.*, expression  
30 constructs that place the described genes under the control of a strong promoter), and transgenic animals that express an ICP transgene or "knock-outs" (which can be

conditional) that do not express a functional ICP (see, for example, PCT Applic. No. PCT/US98/03243, filed February 20, 1998, herein incorporated by reference). In addition to knock-outs, an additional aspect of the present invention includes animals having genetically engineered mutations in at least one of the described genes that modify the  
5 activity or expression of the ICP (*i.e.*, point mutations, over-expression mutations, etc.).

Further, the present invention also relates to methods of using the described ICP genes and/or encoded ICPs for the identification of compounds that modulate, *i.e.*, act as agonists or antagonists, of ICP gene expression and/or ICP activity. Such compounds can be used as therapeutic agents for the treatment of various symptomatic representations of  
10 biological disorders or imbalances.

#### **4. DESCRIPTION OF THE SEQUENCE LISTING**

The Sequence Listing provides the sequence of the described ICP polynucleotides, and the amino acid sequences encoded thereby.

#### **5. DETAILED DESCRIPTION OF THE INVENTION**

15 The human ICPs described for the first time herein, are novel transporter proteins that are expressed by genes present in human cells. Given their similarity to ion channel proteins, the described ICPs are presumably involved in the translocation of ions across lipid bilayers (*i.e.*, membranes). Ion channel proteins are directly or indirectly involved in a wide variety of roles in the body including, but not limited to, recycling neurotransmitters,  
20 control of blood volume/pressure, molecular transport, nutrient and fluid absorption, energy production, etc. Interfering with, neutralizing, or enhancing ICP function can thus effect a wide variety of physiological changes. Because of their biological significance, ion channel proteins have been subjected to substantial scientific/commercial scrutiny (see for example U.S. Patent No. 5,380,836 which is herein incorporated by reference in its entirety).

25 The invention encompasses the use of the described ICP gene nucleotides, ICPs, peptides and fusions derived therefrom, as well as antibodies, preferably humanized monoclonal antibodies, or binding fragments, domains, or fusion proteins thereof, to the ICPs (which can, for example, act as ICP agonists or antagonists), antagonists that inhibit ICP function or expression, or agonists that activate ICP activity or increase ICP expression,

or can be used in the diagnosis and treatment of ICP-related diseases or disorders.

Examples of such ICP-related diseases include, but are not limited to, seizures, mental illness, dementia, Alzheimer's disease, depression, kidney disease, high or low blood pressure, cardiopulmonary disease, side-effects of infectious diseases, infertility, and

5 arrhythmia.

In particular, the invention described in the subsections below encompasses ICPs, ICP polypeptides or peptides corresponding to functional domains of the ICPs (*e.g.*, ECD, TM or CD), mutated, truncated or deleted forms of the ICPs (*e.g.*, modified versions missing one or more functional domains or portions thereof, such as,  $\Delta$ ECD,  $\Delta$ TM and/or

10  $\Delta$ CD), ICP fusion proteins (*e.g.*, an ICP or a functional domain of an ICP, such as an ECD, fused to an unrelated protein or peptide such as an immunoglobulin constant region, *i.e.*, IgFc), nucleotide sequences encoding such products, and host cell expression systems that can produce such ICP products.

The invention also encompasses antibodies and anti-idiotypic antibodies (including

15 Fab fragments), antagonists and agonists of the described ICPs, as well as compounds or nucleotide constructs that inhibit the expression of the ICP genes (transcription factor inhibitors, antisense and ribozyme molecules, or gene or regulatory sequence replacement constructs), or promote expression of ICPs (*e.g.*, expression constructs in which ICP coding sequences are operatively associated with expression control elements such as promoters,

20 promoter/enhancers, etc.). The invention also relates to host cells and animals genetically engineered to express the human ICPs (or mutant variants thereof) or to inhibit or "knock-out" expression of an animal's endogenous ICP gene. Another variation of the such knock-out animals includes "knock-in" animals where the endogenous copy of the animal gene has been replaced by related activity encoded by a sequence that is not native to the host animal

25 (*e.g.*, where a human ortholog has been used to replace the corresponding endogenous gene).

The described ICPs, or peptides therefrom, ICP fusion proteins, ICP nucleotide sequences, antibodies, antagonists and agonists can be useful for the detection of mutant ICPs or inappropriately expressed variants of the ICP for the diagnosis of ICP-related

30 diseases or disorders. The ICPs, or peptides therefrom, ICP fusion proteins, ICP gene nucleotide sequences, host cell expression systems, antibodies, antagonists, agonists and

genetically engineered cells and animals can also be used for screening for drugs (or high throughput screening of compound "libraries") effective in the treatment of the symptomatic or phenotypic manifestations of perturbing the normal function of a ICP in the body. The use of engineered host cells and/or animals may offer an advantage in that such systems  
5 allow not only for the identification of compounds that bind to an ECD of an ICP, but can also identify compounds that affect the activity of the ICP.

Finally, the ICP products (especially soluble derivatives such as peptides corresponding to the ICP ECD, or truncated polypeptides lacking one or more TM domains) and fusion protein products (especially ICP-Ig fusion proteins, *i.e.*, fusions of an ICP, or a  
10 domain of an ICP, *e.g.*, ECD,  $\Delta$ TM to an IgFc), antibodies and anti-idiotypic antibodies (including Fab fragments), antagonists, or agonists can be used in the treatment of disease. For example, the administration of an effective amount of soluble ICP ECD,  $\Delta$ TM, or an ECD-IgFc fusion protein or an anti-idiotypic antibody (or its Fab) that mimics the ICP ECD would "mop up" or "neutralize" ICP effector, modulatory, or activator, ligands, and prevent  
15 or reduce ion channel expression and/or activity. Alternatively, such ICP derivatives could directly associate with the corresponding ion channel macromolecule and effect ion channel function by, for example, competing with normal ICP function/assembly.

Nucleotide constructs encoding such ICP products can be used to genetically engineer host cells to express such products *in vivo*; these genetically engineered cells  
20 function as "bioreactors" in the body delivering a continuous supply of an ICP, a ICP peptide, soluble ECD or  $\Delta$ TM or an ICP fusion protein that can either enhance ICP activity or inhibit ICP activity. Nucleotide constructs encoding a functional ICP, mutant ICP variants, as well as antisense and ribozyme molecules can thus be used in "gene therapy" approaches for the modulation of ICP expression, or associated ion channel activity. Thus,  
25 the invention also encompasses pharmaceutical formulations and methods for treating biological disorders that comprise one or more of the described ICP genes or products. Suitable vectors/systems for such gene therapy or gene delivery applications include, but are not limited to, retrovirus, lentivirus, SIV, HIV, adenovirus, adeno-associated virus, lipid associated (and particularly cationic lipid associated) polynucleotide preparations, micro  
30 carrier beads or lattices, herpes virus vectors, hepatitis virus vectors, polynucleotide-containing emulsions, formulations containing "naked" DNA, etc.

Various aspects of the invention are described in greater detail in the subsections below.

### 5.1 THE ICP GENES

The cDNA sequences and deduced amino acid sequences of the presently described human ICPs are presented in the Sequence Listing. SEQ ID NOS: 1-4 describe variants of a novel human ion channel protein that is almost ubiquitously expressed when assayed by PCR including, *inter alia*, brain, heart, kidney, bone, and testis. Northern analysis reveals a major transcript of approximately 1.9-2.1 kb that is predominantly expressed in testis. cDNAs used for sequencing were isolated from a testis cDNA library. Given the pattern of expression, the ICPs, or drugs targeting the same may also be useful in the treatment of infertility or impotence, as birth control, or to regulate androgen (*i.e.*, testosterone, etc.) production,

Homology studies using SEQ ID NOS: 1-4 indicated that the described molecules share substantial similarity with, *inter alia*, sodium channel protein alpha subunits from a variety of mammalian sources and the Drosophila paralytic protein. SEQ ID NOS:3 and 4 describe an alternative ICP sequence.

The ICPs of the present invention include: (a) the human DNA sequences encoding the ICPs presented in the Sequence Listing and additionally contemplate any nucleotide sequences encoding a contiguous and functional ICP open reading frame (ORF) that hybridizes to a complement of a DNA sequence presented in the Sequence Listing under highly stringent conditions, *e.g.*, hybridization to filter-bound DNA in 0.5 M NaHPO<sub>4</sub>, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1xSSC/0.1% SDS at 68°C (Ausubel F.M. *et al.*, eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York, at p. 2.10.3) and encodes a functionally equivalent gene product. Additionally contemplated are any nucleotide sequences that hybridize to the complement of a DNA sequence that encodes and expresses an amino acid sequence presented in the Sequence Listing under moderately stringent conditions, *e.g.*, washing in 0.2xSSC/0.1% SDS at 42°C (Ausubel *et al.*, 1989, *supra*), yet which still encode a functionally equivalent ICP product. Functional equivalents of ICPs include naturally occurring ICPs present in other species, and mutant ICPs whether



naturally occurring or engineered. The invention also includes degenerate variants of the disclosed sequence.

The invention also includes nucleic acid molecules, preferably DNA molecules, that hybridize to, and are therefore the complements of, the described ICP gene sequences. Such hybridization conditions may be highly stringent or less highly stringent, as described above. In instances wherein the nucleic acid molecules are deoxyoligonucleotides ("DNA oligos"), such molecules are typically about 16 to about 100 bases long, about 20 to about 80, or about 34 to about 45 bases long, or any variation or combination of sizes represented therein and which incorporate a contiguous region of sequence first disclosed in the Sequence Listing. The described oligonucleotides can be used in conjunction with the polymerase chain reaction (PCR) to screen libraries, isolate clones, and prepare cloning and sequencing templates, etc.. Alternatively, the oligonucleotides can be used as hybridization probes. For oligonucleotide probes, highly stringent conditions may refer, *e.g.*, to washing in 6xSSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48° C (for 17-base oligos), 55° C (for 20-base oligos), and 60° C (for 23-base oligos). These nucleic acid molecules, or polynucleotides including such sequences, may encode or act as ICP antisense molecules useful, for example, in ICP gene regulation (and/or as antisense primers in amplification reactions of ICP gene nucleic acid sequences). With respect to ICP gene regulation, such techniques can be used to regulate the biological functions affected by the described ICPs. Further, such sequences may be used as part of ribozyme and/or triple helix sequences that are also useful for ICP gene regulation.

Low stringency conditions are well known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see, for example, Sambrook *et al.*, 1989, *Molecular Cloning, A Laboratory Manual* (and periodic updates thereof), Cold Springs Harbor Press, N.Y.; and Ausubel *et al.*, 1989, *Current Protocols in Molecular Biology*, Green Publishing Associates and Wiley Interscience, N.Y.

Alternatively, suitably labeled ICP nucleotide probes can be used to screen a human genomic library using appropriately stringent conditions, or using PCR. The identification and characterization of human genomic clones is helpful for identifying polymorphisms, determining the genomic structure of a given locus/allele, and designing diagnostic tests.

For example, sequences derived from regions adjacent to the intron/exon boundaries of the human gene can be used to design primers for use in amplification assays to detect mutations within the exons, introns, splice sites (*e.g.*, splice acceptor and/or donor sites), etc., that can be used in diagnostics and pharmacogenomics.

5 Further, an ICP gene homolog can be isolated from nucleic acid from the organism of interest by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of amino acid sequences present within an ICP product disclosed herein. The template for the reaction may be total RNA, mRNA, and/or cDNA obtained by reverse transcription of mRNA prepared from, for example, human or non-human cell lines or  
10 tissue that are known or suspected to express an ICP gene.

The PCR product can be subcloned and sequenced to ensure that the amplified sequences represent the sequence of the desired ICP gene. The PCR fragment may then be used to isolate a full length cDNA clone by a variety of methods. For example, the amplified fragment can be labeled and used to screen a cDNA library such as a  
15 bacteriophage cDNA library. Alternatively, the labeled fragment can be used to isolate genomic clones via the screening of a genomic library.

PCR technology can also be utilized to isolate full length cDNA sequences. For example, RNA may be isolated, using standard procedures, from an appropriate cellular or tissue source (*i.e.*, one known, or suspected, to express an ICP gene, such as, for example,  
20 testis tissue). A reverse transcription (RT) reaction can be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" using a standard terminal transferase reaction, the hybrid may be digested with RNase H, and second strand synthesis may then be primed with a complementary primer. Thus, cDNA  
25 sequences upstream of the amplified fragment can easily be isolated. For a review of cloning strategies which may be used, see *e.g.*, Sambrook *et al.*, 1989, *supra*.

A cDNA of a mutant ICP gene may be isolated, for example, by using PCR. In this case, the first cDNA strand can be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known or suspected to be expressed in an individual  
30 putatively carrying a mutant ICP encoding allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an

oligonucleotide that specifically hybridizes to the 5' end of the normal gene. Using these two primers, the product is then amplified via PCR, optionally cloned into a suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the mutant ICP allele to that of a  
5 corresponding normal ICP allele, the mutation(s) responsible for the loss or alteration of function of the mutant ICP product can be ascertained.

Alternatively, a genomic library can be constructed using DNA obtained from an individual suspected of or known to carry a mutant ICP allele, or a cDNA library can be constructed using RNA from a tissue known, or suspected, to express a mutant or  
10 alternatively spliced ICP allele. A normal ICP gene, or any suitable fragment thereof, can then be labeled and used as a probe to identify the corresponding mutant ICP allele in such libraries. Clones containing the mutant ICP gene sequences can then be purified and subjected to sequence analysis according to methods well known to those of skill in the art.

Additionally, an expression library can be constructed utilizing cDNA synthesized  
15 from, for example, RNA isolated from a tissue known, or suspected, to express a mutant ICP allele in an individual suspected of or known to carry such a mutant allele. In this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against a normal ICP product, as described, below, in Section 5.3. (For screening techniques, see, for  
20 example, Harlow, E. and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor.)

Additionally, screening can be accomplished using labeled ICP fusion proteins, such as, for example, AP-ICP or ICP-AP fusion proteins. In cases where an ICP mutation results in an expressed gene product with altered function (*e.g.*, as a result of a missense or a  
25 frameshift mutation), a polyclonal set of antibodies to a given ICP are likely to cross-react with the corresponding mutant ICP gene product. Library clones detected by their reaction with such labeled antibodies can be purified and subjected to sequence analysis according to methods well known to those of skill in the art.

The invention also encompasses nucleotide sequences that encode mutant ICPs,  
30 peptide fragments of a ICP, truncated ICPs, and ICP fusion proteins. These include, but are not limited to nucleotide sequences encoding mutant ICPs described in section 5.2 *infra*;

polypeptides or peptides corresponding to one or more ECD, TM and/or CD domains of an ICP or any portions of such domains; truncated ICPs in which one or more of the domains are deleted, *e.g.*, a soluble ICP lacking TM or both the TM and CD regions, or a truncated, nonfunctional ICP lacking all or a portion of, for example, a CD region. Nucleotides  
5 encoding fusion proteins may include, but are not limited to, full length ICP sequences, truncated ICPs, or nucleotides encoding peptide fragments of an ICP fused to an unrelated protein or peptide, such as for example, a transmembrane sequence, which anchors an ICP ECD to the cell membrane; an Ig Fc domain which increases the stability and half life of the resulting fusion protein (*e.g.*, ICP-Ig) in the bloodstream; or an enzyme, fluorescent protein,  
10 luminescent protein which can be used as a marker.

The invention also encompasses (a) DNA vectors that contain any of the ICP coding sequence and/or the complements thereof (*i.e.*, antisense); (b) DNA expression vectors that contain any portion of an ICP coding sequence operatively associated with a regulatory element that directs the expression of the coding sequence; and (c) genetically engineered  
15 host cells engineered to contain ICP coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell. As used herein, regulatory elements include, but are not limited to, inducible and non-inducible promoters, enhancers, operators and other elements known to those skilled in the art that drive and regulate expression. Such regulatory elements include, but are not limited to, the  
20 cytomegalovirus hCMV immediate early gene, regulatable, viral (particularly retroviral LTR promoters) the early or late promoters of SV40 adenovirus, the *lac* system, the *trp* system, the *tac* system, the *trc* system, the major operator and promoter regions of phage lambda, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase (PGK), the promoters of acid phosphatase, and the promoters of the yeast  $\alpha$ -mating factors.

25

## **5.2 ICP PRODUCTS**

ICPs, peptide fragments therefrom, mutated, truncated or deleted forms of ICPs and/or ICP fusion proteins can be prepared for a variety of uses, including but not limited to the generation of antibodies, as reagents in diagnostic assays, the identification of other cellular gene products related to, or that interact with, an ICP, as reagents in assays for

screening for compounds that can be as pharmaceutical reagents useful in the therapeutic treatment of mental, biological, or medical disorders and diseases.

The Sequence Listing discloses the amino acid sequences encoded by the described ICP genes. The described ICPs have an initiator methionine in a DNA sequence context  
5 consistent with a translation initiation site, followed by a initiator codon.

The ICP sequences of the present invention include the nucleotide and amino acid sequences presented in the Sequence Listing as well as analogues and derivatives thereof. Further, corresponding ICP homologues from other species are encompassed by the invention. In fact, any ICP protein encoded by the nucleotide sequences of the ICP genes  
10 described in Section 5.1, above, are within the scope of the invention, as are any novel polynucleotide sequences encoding all or any novel portion of an amino acid sequence presented in the Sequence Listing. The degenerate nature of the genetic code is well known, and, accordingly, each amino acid presented in the Sequence Listing, is generically representative of the well known nucleic acid "triplet" codon, or in many cases codons, that  
15 can encode the amino acid. As such, as contemplated herein, an amino acid sequence presented in the Sequence Listing, when taken together with the genetic code (see, for example, Table 4-1 at page 109 of "Molecular Cell Biology", 1986, J. Darnell *et al.* eds., Scientific American Books, New York, NY, herein incorporated by reference) is generically representative of all the various permutations and combinations of nucleic acid sequences  
20 that can encode such amino acid sequence (as well as such variants that are biased by human codon usage frequency tables).

The invention also encompasses proteins that are functionally equivalent to an ICP encoded by a nucleotide sequence described in Section 5.1, as judged by any of a number of criteria, including but not limited to the ability to bind or transport a ligand of the ICP, the  
25 ability to effect an identical or complementary biological pathway, a change in cellular metabolism (*e.g.*, ion flux, tyrosine phosphorylation, etc.), or to effect the same change in phenotype when the ICP equivalent is present in an appropriate cell type (such as the amelioration, prevention or delay of a biochemical, biophysical, or overt phenotype). Such functionally equivalent ICPs include but are not limited to additions or substitutions of  
30 amino acid residues within the amino acid sequence encoded by a ICP gene sequence described above, in Section 5.1, but which result in a silent change, thus producing a

functionally equivalent gene product. Amino acid substitutions can be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

While random mutations can be made to ICP gene DNA (using random mutagenesis techniques well known to those skilled in the art) and the resulting mutant ICPs tested for activity, site-directed mutations of an ICP coding sequence can be engineered (using site-directed mutagenesis techniques well known to those skilled in the art) to generate mutant ICPs with increased function, *e.g.*, higher binding/transport affinity for the transporter substrate, or decreased function. One starting point for such analysis is by aligning the disclosed human sequences with corresponding gene/protein sequences from, for example, other mammals in order to identify amino acid sequences and motifs that are conserved between different species. Non-conservative changes can be engineered at variable positions to alter function, signal transduction capability, or both. Alternatively, where alteration of function is desired, deletion or non-conservative alterations of the conserved regions (*i.e.*, identical amino acids) can be engineered. For example, deletion or non-conservative alterations (substitutions or insertions) of the various conserved transmembrane domains.

Other mutations to an ICP coding sequence can be made to generate ICPs that are better suited for expression, scale up, etc. in the host cells chosen. For example, cysteine residues can be deleted or substituted with another amino acid in order to eliminate disulfide bridges; N-linked glycosylation sites can be altered or eliminated to achieve, for example, expression of a homogeneous product that is more easily recovered and purified from yeast hosts which are known to hyperglycosylate N-linked sites. To this end, a variety of amino acid substitutions at one or both of the first or third amino acid positions of any one or more of the glycosylation recognition sequences which occur in an ECD (N-X-S or N-X-T), and/or an amino acid deletion at the second position of any one or more such recognition

sequences in an ECD will prevent glycosylation of the ICP at the modified tripeptide sequence. (See, *e.g.*, Miyajima *et al.*, 1986, EMBO J. 5(6):1193-1197).

Peptides corresponding to one or more domains of an ICP (*e.g.*, ECD, TM, CD, etc.), truncated or deleted ICPs (*e.g.*, an ICP in which an ECD, TM and/or CD, or any portion thereof, is deleted) as well as fusion proteins in which a full length ICP, ICP peptide, or truncated ICP is fused to an unrelated protein, are also within the scope of the invention, and can be designed on the basis of the presently disclosed ICP gene nucleotide and ICP amino acid sequences. Such fusion proteins include but are not limited to IgFc fusions which stabilize an ICP or ICP peptide and prolong half-life *in vivo*; or fusions to any amino acid sequence that allows the fusion protein to be anchored to the cell membrane, allowing an ECD to be exhibited on the cell surface; or fusions to an enzyme, fluorescent protein, or luminescent protein which provide a marker function.

While an ICP, and corresponding peptides, can be chemically synthesized (*e.g.*, see Creighton, 1983, Proteins: Structures and Molecular Principles, W.H. Freeman & Co., N.Y.), large polypeptides derived from an ICP and especially a full-length ICP product can be advantageously produced by recombinant DNA technology using techniques well known in the art for expressing proteins. Such methods can be used to construct expression vectors containing an ICP gene nucleotide sequence described in Section 5.1 and appropriate transcriptional and translational control signals. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. See, for example, the techniques described in Sambrook *et al.*, 1989, *supra*, and Ausubel *et al.*, 1989, *supra*. Alternatively, RNA corresponding to all or a portion of a transcript encoded by an ICP gene sequence may be chemically synthesized using, for example, synthesizers. See, for example, the techniques described in "Oligonucleotide Synthesis", 1984, Gait, M.J. ed., IRL Press, Oxford, which is incorporated by reference herein in its entirety.

A variety of host-expression vector systems can be utilized to express an ICP gene nucleotide sequence of the invention. Where the ICP peptide or polypeptide is a soluble derivative (*e.g.*, ICP peptides corresponding to an ECD; truncated or deleted ICP in which a TM and/or CD are deleted) the peptide or polypeptide can be recovered from the culture, *i.e.*, from the host cell in cases where the ICP peptide or polypeptide is not secreted, and

from the culture media in cases where the ICP peptide or polypeptide is secreted by the cells. However, such expression systems also encompass engineered host cells that express an ICP, or a functional equivalent, *in situ*, *i.e.*, anchored in the cell membrane. Purification or enrichment of an ICP from such expression systems can be accomplished using

5 appropriate detergents and lipid micelles, and methods well known to those skilled in the art. However, such engineered host cells themselves may be used in situations where it is important not only to retain the structural and functional characteristics of a ICP, but to assess biological activity, *e.g.*, in drug screening assays.

The expression systems that can be used for purposes of the invention include but  
10 are not limited to microorganisms such as bacteria (*e.g.*, *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing an ICP nucleotide sequence; yeast (*e.g.*, *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing an ICP nucleotide sequence; insect cell systems infected with recombinant virus expression vectors (*e.g.*, baculo virus)  
15 containing an ICP sequence; plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid) containing a ICP nucleotide sequence; or mammalian cell systems (*e.g.*, COS, CHO, BHK, 293, 3T3, etc.) harboring recombinant expression constructs containing promoters derived from the genome of  
20 mammalian cells (*e.g.*, metallothionein promoter) or from mammalian viruses (*e.g.*, the adenovirus late promoter; the vaccinia virus 7.5K promoter, etc.).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the ICP product to be expressed. For example, when a large quantity of such a protein is to be produced for the generation of pharmaceutical  
25 compositions comprising an ICP, or for raising antibodies to an ICP, or corresponding peptide, for example, vectors that direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther *et al.*, 1983, EMBO J. 2:1791), in which an ICP gene coding sequence may be ligated into the vector in frame with the *lacZ*  
30 coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke & Schuster, 1989, J. Biol. Chem. 264:5503-



5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The PGEX vectors are  
5 designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In an insect system, *Autographa californica* nuclear polyhydrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. An ICP gene coding sequence may be cloned individually into non-essential regions (for  
10 example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of an ICP gene coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in  
15 which the inserted gene is expressed (*e.g.*, see Smith *et al.*, 1983, J. Virol. 46: 584; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, a nucleotide sequence from an ICP gene can be ligated to an adenovirus transcription/translation control  
20 complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric gene can then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing an ICP product in infected host cells (*e.g.*, See Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:3655-3659). Specific initiation  
25 signals can also be required for efficient translation of inserted ICP gene nucleotide sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire ICP gene or cDNA, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of an ICP coding  
30 sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with

the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can have a variety of origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (See

5 Bittner *et al.*, 1987, *Methods in Enzymol.* 153:516-544).

In addition, a host cell strain can be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (*e.g.*, glycosylation) and processing (*e.g.*, cleavage) of protein products may be important for the function of the protein. Different host cells have charac-  
10 teristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used.  
15 Such mammalian host cells include, but are not limited to, CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, and WI38 cell lines.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express an ICP can be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be  
20 transformed with DNA controlled by appropriate expression control elements (*e.g.*, promoter, enhancer sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection  
25 and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines that express an ICP product. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of an ICP.

30 A number of selection systems can be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, *et al.*, 1977, *Cell* 11:223), hypoxanthine-guanine

phosphoribosyltransferase (Szybalski & Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48:2026), and adenine phosphoribosyltransferase (Lowy, *et al.*, 1980, Cell 22:817) genes can be employed in corresponding *tk<sup>-</sup>*, *hgp<sup>r</sup>* or *ap<sup>r</sup>* cells. Also, antimetabolite resistance can be used as the basis of selection for the following genes: *dhfr*, which confers resistance to methotrexate (Wigler, *et al.*, 1980, Natl. Acad. Sci. USA 77:3567; O'Hare, *et al.*, 1981, Proc. Natl. Acad. Sci. USA 78:1527); *gpt*, which confers resistance to mycophenolic acid (Mulligan & Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); *neo*, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, *et al.*, 1981, J. Mol. Biol. 150:1); and *hygro*, which confers resistance to hygromycin (Santerre, *et al.*, 1984, Gene 30:147).

10 Alternatively, any fusion protein may be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht *et al.* allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, *et al.*, 1991, Proc. Natl. Acad. Sci. USA 88: 8972-8976). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that  
15 the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni<sup>2+</sup>-nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

ICP product(s) can also be expressed in transgenic animals. Animals of any species,  
20 including, but not limited to, worms, mice, rats, rabbits, guinea pigs, pigs, micro-pigs, birds, goats, dogs, cats, and non-human primates, *e.g.*, baboons, monkeys, and chimpanzees may be used to generate ICP transgenic animals.

Any technique known in the art may be used to introduce an ICP transgene into animals to produce the founder lines of transgenic animals. Such techniques include, but are  
25 not limited to pronuclear microinjection (Hoppe, P.C. and Wagner, T.E., 1989, U.S. Pat. No. 4,873,191); retrovirus mediated gene transfer into germ line cells (Van der Putten *et al.*, 1985, Proc. Natl. Acad. Sci., USA 82:6148-6152); gene targeting in embryonic stem cells (Thompson *et al.*, 1989, Cell 56:313-321); electroporation of embryos (Lo, 1983, Mol Cell Biol. 3:1803-1814); and sperm-mediated gene transfer (Lavitrano *et al.*, 1989, Cell 57:717-  
30 723); etc. For a review of such techniques, see Gordon, 1989, Transgenic Animals, Intl. Rev. Cytol. 115:171-229, which is incorporated by reference herein in its entirety.

The present invention provides for transgenic animals that carry an ICP transgene in all of their cells, as well as animals which carry the transgene in some, but not all their cells, *i.e.*, mosaic animals or somatic cell transgenic animals. The transgene may be integrated as a single transgene or in concatamers, *e.g.*, head-to-head tandems or head-to-tail tandems.

5 The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko *et al.*, 1992, Proc. Natl. Acad. Sci. USA 89:6232-6236. The regulatory sequences required for such cell type-specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

10 When it is desired that the ICP transgene be integrated into the chromosomal site of the endogenous ICP gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to an endogenous ICP gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, a mutagenic sequence into the targeted ICP gene that effectively  
15 disrupts the function of the endogenous gene (*i.e.*, "knockout" cells and animals).

The transgene may also be selectively introduced into a particular cell type, thus inactivating an endogenous ICP gene in only that cell type, by following, for example, the teaching of Gu *et al.*, 1994, Science, 265:103-106. The regulatory sequences required for such cell type specific inactivation will depend upon the particular cell type of interest, and  
20 will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of a recombinant ICP gene can be assayed utilizing standard techniques. Initial screening can be accomplished by Southern blot analysis or PCR to analyze animal tissues to determine whether the transgene has integrated into the genome. The level of mRNA expression by the transgene in the  
25 tissues of the transgenic animals can also be assessed using techniques which include but are not limited to Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and RT-PCR. Samples of tissue that express an ICP gene can also be evaluated using, for example, immunocytochemical methods using antibodies specific for the ICP product of the transgene.

### 5.3 ANTIBODIES TO ICPs

Antibodies that specifically recognize one or more epitopes of an ICP, or epitopes of conserved variants of an ICP, or peptide fragments of an ICP are also encompassed by the invention. Such antibodies include but are not limited to polyclonal antibodies, monoclonal  
5 antibodies (mAbs), human, humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')<sub>2</sub> fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above.

The antibodies of the invention can be used, for example, in the detection of an ICP in a biological sample and can therefore be utilized as part of a diagnostic or prognostic  
10 technique whereby patients are tested for abnormal amounts of ICP expression or activity. Such antibodies may also be utilized in conjunction with, for example, compound screening schemes, as described, below, in Section 5.5, for the evaluation of the effect of test compounds on the expression and/or activity of an ICP product. Additionally, such antibodies can be used in conjunction with gene therapy to, for example, evaluate the  
15 expression of normal and/or an engineered ICP by cells prior to their introduction into the patient. Such antibodies may additionally be used as a method for the inhibition of abnormal ICP activity. Thus, such antibodies can be utilized as part of treatment methods for ICP-involved biological disorders.

For the production of antibodies, various host animals can be immunized by injection  
20 with the ICP, an ICP peptide (*e.g.*, one corresponding to a functional domain of the receptor, such as an ECD, TM or CD), truncated ICP polypeptides (ICPs in which one or more domains, *e.g.*, a TM or CD has been deleted, or a portion thereof), functional equivalents of an ICP, or ICP mutants. Such host animals may include, but are not limited to, rabbits, goats, mice, and rats, to name but a few. Various adjuvants can be used to  
25 increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*. Polyclonal  
30 antibodies are heterogeneous populations of antibody molecules derived from the sera of the immunized animals.

Monoclonal antibodies are homogeneous populations of antibodies that bind a particular antigen, and can be obtained by any technique which provides for the production of antibody molecules by, for example, cell lines in culture. These include, but are not limited to, the hybridoma technique of Kohler and Milstein, (1975, *Nature* 256:495-497; and  
5 U.S. Patent No. 4,376,110), the human B-cell hybridoma technique (Kosbor *et al.*, 1983, *Immunology Today* 4:72; Cole *et al.*, 1983, *Proc. Natl. Acad. Sci. USA* 80:2026-2030), and the EBV-hybridoma technique (Cole *et al.*, 1985, *Monoclonal Antibodies And Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma  
10 producing the mAb of this invention can be cultivated in vitro or in vivo. Production of high titers of mAbs in vivo makes this the presently preferred method of production.

Additionally, recombinant antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. A chimeric  
15 antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region. (See, *e.g.*, Cabilly *et al.*, U.S. Patent No. 4,816,567; and Boss *et al.*, U.S. Patent No. 4,816,397, which are incorporated herein by reference in their entirety.) Humanized antibodies are antibody molecules from non-human species having  
20 one or more complementarily determining regions (CDRs) from the non-human species and a framework region from a human immunoglobulin molecule. (See, *e.g.*, Queen, U.S. Patent No. 5,585,089, which is incorporated herein by reference in its entirety.) Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT Publication No. WO 87/02671;  
25 European Patent Application 184,187; European Patent Application 171,496; European Patent Application 173,494; PCT Publication No. WO 86/01533; U.S. Patent No. 4,816,567; European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559; Morrison (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *Bio/Techniques* 4:214; U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature*

321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced, for example, using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. The transgenic mice are immunized in the normal fashion with a selected antigen, *e.g.*, all or a portion of a polypeptide of the invention. Monoclonal antibodies directed against the antigen can be obtained using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA and IgE antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar (1995, *Int. Rev. Immunol.* 13:65-93). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, *see, e.g.*, U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806. In addition, companies such as Abgenix, Inc. (Fremont, CA), can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, *e.g.*, a mouse antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jespers *et al.* (1994) *Bio/technology* 12:899-903).

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778; Bird, 1988, *Science* 242:423-426; Huston *et al.*, 1988, *Proc. Natl. Acad. Sci. USA* 85:5879-5883; and Ward *et al.*, 1989, *Nature* 334:544-546) can be adapted to produce single chain antibodies against ICP products. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, such fragments include but are not limited to: the F(ab')<sub>2</sub>

fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed (Huse *et al.*, 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the  
5 desired specificity.

Antibodies to an ICP can be utilized to generate anti-idiotypic antibodies that "mimic" an ICP, using techniques well known to those skilled in the art. (See, *e.g.*, Greenspan & Bona, 1993, FASEB J 7(5):437-444; and Nissinoff, 1991, J. Immunol. 147(8):2429-2438). For example, antibodies that bind to an ICP ECD and competitively  
10 inhibit the binding of a ligand or accessory molecule of an ICP can be used to generate anti-idiotypes that "mimic" an ICP ECD and, therefore, bind and neutralize a ligand or a, ICP accessory molecule. Such neutralizing anti-idiotypes or Fab fragments of such anti-idiotypic antibodies can be used in therapeutic regimens involving the regulation of ICP activity..

#### 5.4 DIAGNOSIS OF ABNORMALITIES RELATED TO AN ICP

15 A variety of methods can be employed for the diagnostic and prognostic evaluation of disorders related to ICP function, and for the identification of subjects having a predisposition for such disorders.

Such methods can, for example, utilize reagents such as an ICP nucleotide sequence described in Section 5.1, or a portion thereof, or ICP antibodies as described in Section 5.3.  
20 Specifically, such reagents can be used, for example, for: (1) the detection of the presence of ICP gene mutations, or the detection of either the over- or under-expression of ICP mRNA relative to a given phenotype; (2) the detection of either an over- or under-abundance of an ICP relative to a given phenotype; and (3) the detection of perturbations or abnormalities in the transporter function mediate by an ICP.

25 The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one specific ICP nucleotide sequence or ICP antibody reagent described herein, that can be conveniently used, *e.g.*, in clinical settings, to diagnose patients exhibiting biological abnormalities.

For the detection of ICP mutations, any nucleated cell can be used as a starting  
30 source of genomic nucleic acid. For the detection of ICP gene expression or ICP products,



any cell type or tissue in which an ICP gene is expressed, such as, for example, testis cells, can be utilized.

Nucleic acid-based detection techniques are described, below, in Section 5.4.1.

Peptide detection techniques are described, below, in Section 5.4.2.

5

#### **5.4.1 DETECTION OF ICP GENES AND TRANSCRIPTS**

Mutations within an ICP gene can be detected using a number of techniques.

Nucleic acid from any nucleated cell can be used as the starting point for such assay techniques, and can be isolated according to standard nucleic acid preparation procedures that are well known to those of skill in the art.

10

DNA can be used in hybridization or amplification assays of biological samples to detect abnormalities involving ICP gene structure, including point mutations, insertions, deletions and chromosomal rearrangements. Such assays may include, but are not limited to, Southern analyses, single stranded conformational polymorphism analyses (SSCP), and PCR analyses.

15

Such diagnostic methods for the detection of ICP gene specific mutations can involve for example, contacting and incubating nucleic acids including recombinant DNA molecules, cloned genes or degenerate variants thereof, obtained from a sample, *e.g.*, derived from a patient sample or other appropriate cellular source, with one or more labeled nucleic acid reagents including recombinant DNA molecules, cloned genes or degenerate variants thereof, as described in Section 5.1, under conditions favorable for the specific annealing of these reagents to their complementary sequences within an ICP gene.

20

Preferably, the lengths of these nucleic acid reagents are at least 15 to 30 nucleotides. After incubation, all non-annealed nucleic acids are removed from the nucleic acid:ICP gene hybrid. The presence of nucleic acids which have hybridized, if any such molecules exist, is then detected. Using such a detection scheme, the nucleic acid from the cell type or tissue of interest can be immobilized, for example, to a solid support such as a membrane, gene "chip" substrate, or a plastic surface such as that on a microtiter plate or polystyrene beads.

25

In this case, after incubation, non-annealed, labeled nucleic acid reagents of the type described in Section 5.1 are easily removed. Detection of the remaining, annealed, labeled ICP nucleic acid reagents is accomplished using standard techniques well-known to those in

30

the art. The ICP gene sequence(s) to which the nucleic acid reagents have annealed can be compared to the annealing pattern expected from a normal ICP gene sequence in order to determine whether an ICP gene mutation is present.

Alternative diagnostic methods for the detection of ICP gene specific nucleic acid molecules, in patient samples or other appropriate cell sources, may involve their amplification, *e.g.*, by PCR (the experimental embodiment set forth in Mullis, K.B., 1987, U.S. Patent No. 4,683,202), followed by the detection of the amplified molecules using techniques well known to those of skill in the art. The resulting amplified sequences can be compared to those which would be expected if the nucleic acid being amplified contained only normal copies of an ICP gene in order to determine whether an ICP gene mutation exists.

Additionally, well-known genotyping techniques can be performed to identify individuals carrying ICP gene mutations. Such techniques include, for example, the use of restriction fragment length polymorphisms (RFLPs), which involve sequence variations in one of the recognition sites for the specific restriction enzyme used.

Additionally, improved methods for analyzing DNA polymorphisms that can be utilized for the identification of ICP gene mutations have been described that capitalize on the presence of variable numbers of short, tandemly repeated DNA sequences between the restriction enzyme sites. For example, Weber (U.S. Pat. No. 5,075,217, which is incorporated herein by reference in its entirety) describes a DNA marker based on length polymorphisms in blocks of (dC-dA)<sub>n</sub>-(dG-dT)<sub>n</sub> short tandem repeats. The average separation of (dC-dA)<sub>n</sub>-(dG-dT)<sub>n</sub> blocks is estimated to be 30,000-60,000 bp. Markers which are so closely spaced exhibit a high frequency co-inheritance, and are extremely useful in the identification of genetic mutations, such as, for example, mutations within an ICP gene, and the diagnosis of diseases and disorders related to such ICP mutations.

Also, Caskey *et al.* (U.S. Pat. No. 5,364,759, which is incorporated herein by reference in its entirety) describe a DNA profiling assay for detecting short tri and tetra nucleotide repeat sequences. The process includes extracting the DNA of interest, such as an ICP gene, amplifying the extracted DNA, and labeling the repeat sequences to form a genotypic map of the individual's DNA.

The level of ICP gene expression can also be assayed by detecting and measuring ICP transcription. For example, RNA from a cell type or tissue known, or suspected to express an ICP gene, such as testis, may be isolated and tested utilizing hybridization or PCR techniques such as are described, above. The isolated cells can be derived from cell culture or from a patient. The analysis of cells taken from culture may be a necessary step in the assessment of cells to be used as part of a cell-based gene therapy technique or, alternatively, to test the effect of compounds on the expression of an ICP gene. Such analyses may reveal both quantitative and qualitative aspects of the expression pattern of an ICP gene, including activation or inactivation of ICP gene expression.

In one embodiment of such a detection scheme, cDNAs are synthesized from the RNAs of interest (*e.g.*, by reverse transcription of the RNA molecule into cDNA). A sequence within the cDNA is then used as the template for a nucleic acid amplification reaction, such as a PCR amplification reaction, or the like. The nucleic acid reagents used as synthesis initiation reagents (*e.g.*, primers) in the reverse transcription and nucleic acid amplification steps of this method are chosen from among the ICP gene nucleic acid reagents described in Section 5.1. The preferred lengths of such nucleic acid reagents are at least 9-30 nucleotides. For detection of the amplified product, the nucleic acid amplification may be performed using radioactively or non-radioactively labeled nucleotides. Alternatively, enough amplified product may be made such that the product may be visualized by standard ethidium bromide staining, by utilizing any other suitable nucleic acid staining method, or by sequencing.

Additionally, it is possible to perform such ICP gene expression assays "*in situ*", *i.e.*, directly upon tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Nucleic acid reagents such as those described in Section 5.1 may be used as probes and/or primers for such *in situ* procedures (See, for example, Nuovo, G.J., 1992, "PCR *In Situ* Hybridization: Protocols And Applications", Raven Press, NY).

Alternatively, if a sufficient quantity of the appropriate cells can be obtained, standard Northern analysis can be performed to determine the level of mRNA expression of an ICP gene.

### 5.4.2 DETECTION OF ICP PRODUCTS

Antibodies directed against wild type or mutant ICPs or variants or peptide fragments thereof, such as those discussed in Section 5.3, can also be used as diagnostics and prognostics, as described herein. Such diagnostic methods, can be used to detect  
5 abnormalities in the level of ICP gene expression, or abnormalities in the structure and/or temporal, tissue, cellular, or subcellular location of an ICP, and can be performed *in vivo* or *in vitro*, such as, for example, on biopsy tissue.

Additionally, antibodies directed to epitopes of an ICP ECD can be used *in vivo* to detect the pattern and level of expression of an ICP in the body. Such antibodies can be  
10 labeled, *e.g.*, with a radio-opaque or other appropriate compound and injected into a subject in order to visualize binding to an ICP expressed in the body using methods such as X-rays, CAT-scans, or MRI. Labeled antibody fragments, *e.g.*, the Fab or single chain antibody comprising the smallest portion of the antigen binding region, are preferred for this purpose to promote crossing the blood-brain barrier and permit labeling of ICPs expressed in the  
15 brain.

Additionally, any ICP fusion protein or ICP conjugated protein whose presence can be detected, can be administered. For example, ICP fusion or conjugated proteins labeled with a radio-opaque or other appropriate compound can be administered and visualized *in vivo*, as discussed, above for labeled antibodies. Further, such ICP fusion proteins as AP-  
20 ICP on ICP-Ap fusion proteins can be utilized for *in vitro* diagnostic procedures.

Alternatively, immunoassays or fusion protein detection assays, as described above, can be utilized on biopsy and autopsy samples *in vitro* to permit assessment of the expression pattern of an ICP. Such assays are not confined to the use of antibodies that define an ICP ECD, but can include the use of antibodies directed to epitopes of any of the  
25 domains of an ICP, *e.g.*, a ECD, a TM and/or CD. The use of each or all of these labeled antibodies will yield useful information regarding translation and intracellular transport of an ICP to the cell surface, and can identify defects in processing.

The tissue or cell type to be analyzed will generally include those which are known, or suspected, to express an ICP gene, such as, for example, testis cells. The protein isolation  
30 methods employed herein may, for example, be such as those described in Harlow and Lane (Harlow, E. and Lane, D., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor

Laboratory Press, Cold Spring Harbor, New York), which is incorporated herein by reference in its entirety. The isolated cells can be derived from cell culture or from a patient. The analysis of cells taken from culture may be a necessary step in the assessment of cells that could be used as part of a cell-based gene therapy technique or, alternatively, to test the effect of compounds on the expression of an ICP gene.

For example, antibodies, or fragments of antibodies, such as those described, above, in Section 5.3, useful in the present invention may be used to quantitatively or qualitatively detect the presence of ICPs or conserved variants or peptide fragments thereof. This can be accomplished, for example, by immunofluorescence techniques employing a fluorescently labeled antibody (see below, this Section) coupled with light microscopic, flow cytometric, or fluorimetric detection. Such techniques are especially preferred if such ICP gene products are expressed on the cell surface.

The antibodies (or fragments thereof) or ICP fusion or conjugated proteins useful in the present invention may, additionally, be employed histologically, as in immunofluorescence, immunoelectron microscopy or non-immuno assays, for *in situ* detection of ICPs or conserved variants or peptide fragments thereof, or for ICP binding studies (in the case of labeled fusion proteins incorporating ICP accessory proteins).

*In situ* detection may be accomplished by removing a histological specimen from a patient, and applying thereto a labeled antibody or fusion protein of the present invention. The antibody (or fragment) or fusion protein is preferably applied by overlaying the labeled antibody (or fragment) onto a biological sample. Through the use of such a procedure, it is possible to determine not only the presence of an ICP, or conserved variants or peptide fragments, or ICP binding, but also its distribution in the examined tissue. Using the present invention, those of ordinary skill will readily perceive that any of a wide variety of histological methods (such as staining procedures) can be modified in order to achieve such *in situ* detection.

Immunoassays and non-immunoassays for ICPs or conserved variants or peptide fragments thereof will typically comprise incubating a sample, such as a biological fluid, a tissue extract, freshly harvested cells, or lysates of cells which have been incubated in cell culture, in the presence of a detectably labeled antibody capable of identifying an ICP or

conserved variants or peptide fragments thereof, and detecting the bound antibody by any of a number of techniques well-known in the art.

The biological sample may be brought in contact with and immobilized onto a solid phase support or carrier such as nitrocellulose, or other solid support which is capable of  
5 immobilizing cells, cell particles or soluble proteins. The support may then be washed with suitable buffers followed by treatment with the detectably labeled ICP antibody or ICP ligand fusion protein. The solid phase support can then be washed with the buffer a second time to remove unbound antibody or fusion protein. The amount of bound label on solid support may then be detected by conventional means.

10 By "solid phase support or carrier" is intended any support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention. The support material  
15 may have virtually any possible structural configuration so long as the coupled molecule is capable of binding to an antigen or antibody. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads. Those skilled in the art will know many other  
20 suitable carriers for binding antibody or antigen, or will be able to ascertain the same by use of routine experimentation.

The binding activity of a given lot of ICP antibody or ICP ligand fusion protein may be determined according to well known methods. Those skilled in the art will be able to determine operative and optimal assay conditions for each determination by employing  
25 routine experimentation.

With respect to antibodies, one of the ways in which an ICP antibody can be detectably labeled is by linking it to an enzyme that can be used in an enzyme immunoassay (EIA) (Voller, A., "The Enzyme Linked Immunosorbent Assay (ELISA)", 1978, Diagnostic Horizons 2:1-7, Microbiological Associates Quarterly Publication, Walkersville, MD);  
30 Voller, A. *et al.*, 1978, *J. Clin. Pathol.* 31:507-520; Butler, J.E., 1981, *Meth. Enzymol.* 73:482-523; Maggio, E. (ed.), 1980, *Enzyme Immunoassay*, CRC Press, Boca Raton, FL,;

Ishikawa, E. *et al.*, (eds.), 1981, Enzyme Immunoassay, Kigaku Shoin, Tokyo). The enzyme that is bound to the antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety that can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes  
5 that can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase,  
10 glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods which employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Detection can also be accomplished using any of a variety of other immunoassays.  
15 For example, by radioactively labeling the antibodies or antibody fragments, it is possible to detect ICPs through the use of a radioimmunoassay (RIA) (see, for example, Weintraub, B., Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986, which is incorporated by reference herein). The radioactive isotope can be detected by such means as the use of a gamma  
20 counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin,  
25 allophycocyanin, o-phthaldehyde and fluorescamine.

The antibody can also be detectably labeled using fluorescence emitting metals such as  $^{152}\text{Eu}$ , or others of the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

30 The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by

detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, therrromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

Likewise, a bioluminescent compound may be used to label the antibody of the  
5 present invention. Bioluminescence is a type of chemiluminescence found in biological systems in, which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase, and aequorin.

## 10           **5.5 SCREENING ASSAYS FOR COMPOUNDS THAT MODULATE ICP                     EXPRESSION OR ACTIVITY**

The following assays are designed to identify compounds that interact with (*e.g.*, bind to) an ICP (including, but not limited to an ECD or CD of an ICP), compounds that interact with (*e.g.*, bind to) intracellular proteins that interact with an ICP (including but not  
15 limited to the TM and CD of an ICP), compounds that interfere with the interaction of an ICP with transmembrane or intracellular proteins, or such proteins that may be present in cellular organelles, that are associated with ICP-mediated transport, and to compounds that modulate the activity of an ICP gene (*i.e.*, modulate the level of ICP gene expression) or modulate the quantity of an ICP in the cell. Assays may additionally be utilized that identify  
20 compounds that bind to ICP gene regulatory sequences (*e.g.*, promoter sequences) and which may modulate ICP gene expression. See *e.g.*, Platt, K.A., 1994, J. Biol. Chem. 269:28558-28562, which is incorporated herein by reference in its entirety.

The compounds that can be screened in accordance with the invention include but are not limited to peptides, antibodies and fragments thereof, and other organic compounds  
25 (*e.g.*, peptidomimetics) that bind to an ECD, or a corresponding nontransmembrane domain of a organelle or nuclear membrane (in the case where the ICP is associated with an intracellular membrane) of the described ICP and either facilitates or inhibits ICP activity; as well as peptides, antibodies or fragments thereof, and other organic compounds that mimic a domain of an ICP (or a portion thereof) and bind to and "neutralize" ICP accessory proteins.



Such compounds may include, but are not limited to, peptides such as, for example, soluble peptides, including but not limited to members of random peptide libraries; (see, *e.g.*, Lam, K.S. *et al.*, 1991, *Nature* 354:82-84; Houghten, R. *et al.*, 1991, *Nature* 354:84-86), and combinatorial chemistry-derived molecular library made of D- and/or L- configuration amino acids, phosphopeptides (including, but not limited to members of random or partially degenerate, directed phosphopeptide libraries; see, *e.g.*, Songyang, Z. *et al.*, 1993, *Cell* 72:767-778), antibodies (including, but not limited to, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and FAb, F(ab')<sub>2</sub> and FAb expression library fragments, and epitope-binding fragments thereof), and small organic or inorganic molecules.

Other compounds that can be screened in accordance with the invention include but are not limited to small organic molecules that are able to cross the blood-brain barrier, gain entry into an appropriate cell (*e.g.*, in the choroid plexus, the hypothalamus, etc.) and affect the expression of an ICP gene or some other gene involved in an ICP-mediated transport mechanism (*e.g.*, by interacting with the regulatory region or transcription factors involved in ICP gene expression); or such compounds that affect the activity of an ICP (*e.g.*, by inhibiting or enhancing the activity of the ICP) or the activity of some other intracellular factor associated with ICP activity.

Computer modeling and searching technologies permit identification of compounds, or the improvement of already identified compounds, that can modulate ICP expression or activity. Having identified such a compound or composition, the active sites or regions are identified. Such active sites might typically be ligand binding sites. The active site can be identified using methods known in the art including, for example, from the amino acid sequences of peptides, from the nucleotide sequences of nucleic acids, or from study of complexes of the relevant compound or composition with its natural ligand. In the latter case, chemical or X-ray crystallographic methods can be used to find the active site by finding where on the factor the complexed ligand is found.

Next, the three dimensional geometric structure of the active site is determined. This can be done by known methods, including X-ray crystallography, which can determine a complete molecular structure. On the other hand, solid or liquid phase NMR can be used to determine certain intra-molecular distances. Any other experimental method of structure

determination can be used to obtain partial or complete geometric structures. The geometric structures may be measured with a complexed ligand, natural or artificial, which may increase the accuracy of the active site structure determined.

If an incomplete or insufficiently accurate structure is determined, the methods of  
5 computer based numerical modeling can be used to complete the structure or improve its accuracy. Any recognized modeling method may be used, including parameterized models specific to particular biopolymers such as proteins or nucleic acids, molecular dynamics models based on computing molecular motions, statistical mechanics models based on thermal ensembles, or combined models. For most types of models, standard molecular  
10 force fields, representing the forces between constituent atoms and groups, are necessary, and can be selected from force fields known in physical chemistry. The incomplete or less accurate experimental structures can serve as constraints on the complete and more accurate structures computed by these modeling methods.

Finally, having determined the structure of the active site, either experimentally, by  
15 modeling, or by a combination, candidate modulating compounds can be identified by searching databases containing compounds along with information on their molecular structure. Such a search seeks compounds having structures that match the determined active site structure and that interact with the groups defining the active site. Such a search can be manual, but is preferably computer assisted. These compounds found from this search are  
20 potential ICP modulating compounds.

Alternatively, these methods can be used to identify improved modulating compounds from an already known modulating compound or ligand. The composition of the known compound can be modified and the structural effects of modification can be determined using the experimental and computer modeling methods described above applied  
25 to the new composition. The altered structure is then compared to the active site structure of the compound to determine if an improved fit or interaction results. In this manner systematic variations in composition, such as by varying side groups, can be quickly evaluated to obtain modified modulating compounds or ligands of improved specificity or activity.

Further experimental and computer modeling methods useful to identify modulating compounds based upon identification of the active regions of an ICP, and related transport accessory factors will be apparent to those of skill in the art.

Examples of molecular modeling systems are the CHARMM and QUANTA programs (Polygen Corporation, Waltham, MA). CHARMM performs the energy minimization and molecular dynamics functions. QUANTA performs the construction, graphic modeling and analysis of molecular structure. QUANTA allows interactive construction, modification, visualization, and analysis of the behavior of molecules with each other.

10 A number of articles review computer modeling of drugs interactive with specific proteins, such as Rotivinen, *et al.*, 1988, Acta Pharmaceutica Fennica 97:159-166; Ripka, New Scientist 54-57 (June 16, 1988); McKinaly and Rossmann, 1989, Annu. Rev. Pharmacol. Toxicol. 29:111-122; Perry and Davies, OSAR: Quantitative Structure-Activity Relationships in Drug Design pp. 189-193 (Alan R. Liss, Inc. 1989); Lewis and Dean, 1989  
15 Proc. R. Soc. Lond. 236:125-140 and 141-162; and, with respect to a model receptor for nucleic acid components, Askew, *et al.*, 1989, J. Am. Chem. Soc. 111:1082-1090. Other computer programs that screen and graphically depict chemicals are available from companies such as BioDesign, Inc. (Pasadena, CA.), Allelix, Inc. (Mississauga, Ontario, Canada), and Hypercube, Inc. (Cambridge, Ontario). Although these are primarily designed  
20 for application to drugs specific to particular proteins, they can be adapted to design of drugs specific to regions of DNA or RNA, once that region is identified.

Although described above with reference to design and generation of compounds which could alter binding, one could also screen libraries of known compounds, including natural products or synthetic chemicals, and biologically active materials, including proteins,  
25 for compounds which are inhibitors or activators.

Cell-based systems can also be used to identify compounds that bind one of the described ICPs as well as assess the altered activity associated with such binding in living cells. One tool of particular interest for such assays is green fluorescent protein which is described, *inter alia*, in U.S. Patent No. 5,625,048, herein incorporated by reference. Cells  
30 that may be used in such cellular assays include, but are not limited to, leukocytes, or cell lines derived from leukocytes, lymphocytes, stem cells, including embryonic stem cells, and

the like. In addition, expression host cells (*e.g.*, B95 cells, COS cells, CHO cells, OMK cells, fibroblasts, Sf9 cells) genetically engineered to express functional ICP and to respond to activation by the test, or natural, ligand, as measured by a chemical or phenotypic change, or induction of another host cell gene, can be used as an end point in the assay.

#### 5 5.5.1 IN VITRO SCREENING ASSAYS FOR COMPOUNDS THAT BIND TO ICPs

*In vitro* systems may be designed to identify compounds capable of interacting with (*e.g.*, binding to) the described ICPs (including, but not limited to, a ECD or CD of an ICP). Compounds identified may be useful, for example, in modulating the activity of wild type and/or mutant ICP products; may be useful in elaborating the biological function of an ICP; 10 may be utilized in screens for identifying compounds that disrupt normal ICP functions or interactions; or may in themselves disrupt such interactions.

The principle of the assays used to identify compounds that bind to or interact with an ICP involves preparing a reaction mixture of an ICP and a test compound under conditions and for a time sufficient to allow the two components to interact and bind, thus 15 forming a complex which can be removed and/or detected in the reaction mixture. The ICP species used can vary depending upon the goal of the screening assay. For such applications, one can use a full length ICP, or a soluble truncated ICP, *e.g.*, in which the TM and/or CD is deleted from the molecule, a peptide corresponding to a ECD or a fusion protein containing one or more ICP ECD(s) fused to a protein or polypeptide that affords 20 advantages in the assay system (*e.g.*, labeling, isolation of the resulting complex, etc.) can be utilized. Where compounds that interact with the cytoplasmic domain are sought to be identified, peptides corresponding to an ICP CD and fusion proteins containing an ICP CD can be used.

The screening assays can be conducted in a variety of ways. For example, one 25 method to conduct such an assay would involve anchoring an ICP, ICP polypeptide, ICP peptide or fusion protein, or even the test substance onto a solid phase and detecting ICP/test compound complexes anchored on the solid phase at the end of the reaction. In one embodiment of such a method, the ICP reactant can be anchored onto a solid surface, and the test compound, which is not anchored, may be labeled, either directly or indirectly.

In practice, microtiter plates can conveniently be utilized as the solid phase. The anchored component can be immobilized by non-covalent or covalent attachments. Non-covalent attachment can be accomplished by simply coating the solid surface with a solution of the protein and drying. Alternatively, an immobilized antibody, preferably a monoclonal  
5 antibody, specific for the protein to be immobilized may be used to anchor the protein to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the nonimmobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) under conditions such that any complexes  
10 formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the previously nonimmobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously nonimmobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface;  
15 *e.g.*, using a labeled antibody specific for the previously nonimmobilized component (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody).

Alternatively, a reaction can be conducted in a liquid phase, the reaction products separated from unreacted components, and complexes detected; *e.g.*, using an immobilized  
20 antibody specific for an ICP, ICP polypeptide, peptide or fusion protein or the test compound to anchor any complexes formed in solution, and a labeled antibody specific for the other component of the possible complex to detect anchored complexes.

Alternatively, cell-based assays can be used to identify compounds that interact with ICPs. To this end, cell lines that express an ICP, or cell lines (*e.g.*, COS cells, CHO cells,  
25 fibroblasts, etc.) that have been genetically engineered to express an ICP (*e.g.*, by transfection or transduction of ICP gene DNA) can be used. Interaction of the test compound with, for example, an ECD from an ICP expressed by the host cell can be determined by comparison or competition with native ligand.

### 5.5.2 ASSAYS FOR INTRACELLULAR PROTEINS THAT INTERACT WITH ICPs

Any method suitable for detecting protein-protein interactions may be employed for identifying transmembrane proteins or intracellular proteins that interact with the described ICPs. Among the traditional methods which can be employed are co-immunoprecipitation, crosslinking and co-purification through gradients or chromatographic columns of cell lysates or proteins obtained from cell lysates to identify proteins in the lysate that interact with an ICP. For these assays, the ICP component used can be a full length ICP, a soluble derivative lacking the membrane-anchoring region (*e.g.*, a truncated ICP in which a TM is deleted resulting in a truncated molecule containing a ECD fused to a CD), a peptide corresponding to a CD or a fusion protein containing a CD from the described ICP. Once isolated, such an intracellular protein can be identified and can, in turn, be used, in conjunction with standard techniques, to identify proteins with which it interacts. For example, at least a portion of the amino acid sequence of an intracellular protein that interacts with an ICP can be ascertained using techniques well known to those of skill in the art, such as via the Edman degradation technique. (See, *e.g.*, Creighton, 1983, "Proteins: Structures and Molecular Principles", W.H. Freeman & Co., N.Y., pp.34-49). The amino acid sequence obtained can be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for gene sequences encoding such intracellular proteins. Screening may be accomplished, for example, by standard hybridization or PCR techniques. Techniques for the generation of oligonucleotide mixtures and the screening are well-known. (See, *e.g.*, Ausubel, *supra*, and PCR Protocols: A Guide to Methods and Applications, 1990, Innis, M. *et al.*, eds. Academic Press, Inc., New York).

Additionally, methods may be employed which result in the simultaneous identification of genes which encode the transmembrane or intracellular proteins interacting with an ICP. These methods include, for example, probing expression, libraries, in a manner similar to the well known technique of antibody probing of  $\lambda$ gt11 libraries, using labeled ICPs, or a ICP polypeptide, peptide or fusion protein, *e.g.*, an ICP polypeptide or an ICP domain fused to a marker (*e.g.*, an enzyme, fluor, luminescent protein, or dye), or an Ig-Fc domain.

One method that detects protein interactions *in vivo*, the two-hybrid system, is described in detail for illustration only and not by way of limitation. One version of this

system has been described (Chien *et al.*, 1991, Proc. Natl. Acad. Sci. USA, 88:9578-9582) and is commercially available from Clontech (Palo Alto, CA).

Briefly, utilizing such a system, plasmids are constructed that encode two hybrid proteins: one plasmid consists of nucleotides encoding the DNA-binding domain of a transcription activator protein fused to nucleotide sequence encoding an ICP, or an ICP polypeptide, peptide or fusion protein, and the other plasmid includes nucleotides encoding the transcription activator protein's activation domain fused to a cDNA encoding an unknown protein which has been recombined into this plasmid as part of a cDNA library. The DNA-binding domain fusion plasmid and the cDNA library are transformed into a strain of the yeast *Saccharomyces cerevisiae* that contains a reporter gene (*e.g.*, HBS or *lacZ*) whose regulatory region contains the transcription activator's binding site. Either hybrid protein alone cannot activate transcription of the reporter gene: the DNA-binding domain hybrid cannot because it does not provide activation function and the activation domain hybrid cannot because it cannot localize to the activator's binding sites. Interaction of the two hybrid proteins reconstitutes the functional activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product.

The two-hybrid system or related methodology can be used to screen activation domain libraries for proteins that interact with the "bait" gene product. By way of example, and not by way of limitation, an ICP may be used as the bait gene product. Total genomic or cDNA sequences are fused to the DNA encoding an activation domain. This library and a plasmid encoding a hybrid of the bait ICP product fused to the DNA-binding domain are cotransformed into a yeast reporter strain, and the resulting transformants are screened for those that express the reporter gene. For example, and not by way of limitation, a bait ICP gene sequence, such as the open reading frame of the ICP (or a domain of the ICP) can be cloned into a vector such that it is translationally fused to the DNA encoding the DNA-binding domain of the GAL4 protein. These colonies are purified and the library plasmids responsible for reporter gene expression are isolated. DNA sequencing is then used to identify the proteins encoded by the library plasmids.

A cDNA library of the cell line from which proteins that interact with bait ICP product are to be detected can be made using methods routinely practiced in the art. According to the particular system described herein, for example, the cDNA fragments can

be inserted into a vector such that they are translationally fused to the transcriptional activation domain of GAL4. This library can be co-transformed along with the bait ICP gene-GAL4 fusion plasmid into a yeast strain which contains a lacZ gene driven by a promoter which contains GAL4 activation sequence. A cDNA encoded protein, fused to  
5 GAL4 transcriptional activation domain, that interacts with a bait ICP product will reconstitute an active GAL4 protein and thereby drive expression of the HIS3 gene. Colonies that express HIS3 can be detected by their growth on petri dishes containing semi-solid agar based media lacking histidine. The cDNA can then be purified from these strains, and used to produce and isolate the bait ICP gene-interacting protein using techniques  
10 routinely practiced in the art.

### **5.5.3 ASSAYS FOR COMPOUNDS THAT INTERFERE WITH ICP/INTRACELLULAR OR ICP/MEMBRANE MACROMOLECULE INTERACTION**

The macromolecules that interact with an ICP are referred to, for purposes of this  
15 discussion, as "binding partners". These binding partners are likely to be involved in ICP-mediated transport. Therefore, it is desirable to identify compounds that interfere with or disrupt the interaction of such binding partners that can be useful in regulating the activity of ICP and controlling disorders associated with ICP activity.

The basic principle of the assay systems used to identify compounds that interfere  
20 with the interaction between ICP and any binding partner or partners involves preparing a reaction mixture containing an ICP, ICP polypeptide, peptide or fusion protein as described in Sections 5.5.1 and 5.5.2 above, and the binding partner under conditions and for a time sufficient to allow the two to interact and bind, thus forming a complex. In order to test a compound for inhibitory activity, the reaction mixture is prepared in the presence and  
25 absence of the test compound. The test compound may be initially included in the reaction mixture, or may be added at a time subsequent to the addition of the ICP moiety and its binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between the ICP moiety and the binding partner is then detected. The formation of a complex in the control reaction, but not in the reaction  
30 mixture containing the test compound, indicates that the compound interferes with the interaction of the ICP and the interactive binding partner. Additionally, complex formation



within reaction mixtures containing the test compound and normal ICP may also be compared to complex formation within reaction mixtures containing the test compound and a mutant ICP. This comparison may be important in those cases wherein it is desirable to identify compounds that specifically disrupt interactions of mutant, or mutated, ICPs but not  
5 normal ICPs.

The assay for compounds that interfere with the interaction of the described ICP and binding partners can be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either the ICP moiety product or the binding partner onto a solid phase and detecting complexes anchored on the solid phase at the end of  
10 the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction by competition can be identified by conducting the reaction in the presence of the test substance; *i.e.*, by adding the test substance to the reaction mixture prior  
15 to, or simultaneously with, the ICP moiety and interactive binding partner. Alternatively, test compounds that disrupt preformed complexes, *e.g.* compounds with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are described briefly below.

20 In a heterogeneous assay system, either the ICP moiety or an interactive binding partner, is anchored onto a solid surface, while the non-anchored species is labeled, either directly or indirectly. In practice, microtiter plates are conveniently utilized. The anchored species may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished simply by coating the solid surface with a solution of ICP  
25 product or binding partner and drying. Alternatively, an immobilized antibody specific for the species to be anchored may be used to anchor the species to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the partner of the immobilized species is exposed to the coated surface with or without the test compound. After the reaction is complete,  
30 unreacted components are removed (*e.g.*, by washing) and any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid

surface can be accomplished in a number of ways. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; *e.g.*, using a labeled antibody specific for the initially non-immobilized species (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds which inhibit complex formation or which disrupt preformed complexes can be detected.

Alternatively, the reaction can be conducted in a liquid phase in the presence or absence of the test compound, the reaction products separated from unreacted components, and complexes detected; *e.g.*, using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to detect anchored complexes. Again, depending upon the order of addition of reactants to the liquid phase, test compounds which inhibit complex or which disrupt preformed complexes can be identified.

In an alternate embodiment of the invention, a homogeneous assay can be used. In this approach, a preformed complex of an ICP moiety and an interactive binding partner is prepared in which either the ICP or its binding partners is labeled, but the signal generated by the label is quenched due to formation of the complex (see, *e.g.*, U.S. Patent No. 4,109,496 by Rubenstein which utilizes this approach for immunoassays). The addition of a test substance that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances which disrupt ICP/binding partner interaction can be identified.

In a particular embodiment, an ICP fusion can be prepared for immobilization. For example, an ICP or a peptide fragment, *e.g.*, corresponding to a CD, can be fused to a glutathione-S-transferase (GST) gene using a fusion vector, such as pGEX-5X-1, in such a manner that its binding activity is maintained in the resulting fusion protein. The interactive binding partner can be purified and used to raise a monoclonal antibody using methods routinely practiced in the art and described above, in Section 5.3. This antibody can be labeled with the radioactive isotope  $^{125}\text{I}$ , for example, by methods routinely practiced in the art. In a heterogeneous assay, *e.g.*, the GST-ICP fusion protein can be anchored to

glutathione-agarose beads. The interactive binding partner can then be added in the presence or absence of the test compound in a manner that allows interaction and binding to occur. At the end of the reaction period, unbound material can be washed away, and the labeled monoclonal antibody can be added to the system and allowed to bind to the

5 complexed components. The interaction between the ICP product and the interactive binding partner can be detected by measuring the amount of radioactivity that remains associated with the glutathione-agarose beads. A successful inhibition of the interaction by the test compound will result in a decrease in measured radioactivity.

Alternatively, the GST-ICP fusion protein and the interactive binding partner can be

10 mixed together in liquid in the absence of the solid glutathione-agarose beads. The test compound can be added either during or after the species are allowed to interact. This mixture can then be added to the glutathione-agarose beads and unbound material is washed away. Again the extent of inhibition of the ICP/binding partner interaction can be detected by adding the labeled antibody and measuring the radioactivity associated with the beads.

15 In another embodiment of the invention, these same techniques can be employed using peptide fragments that correspond to the binding domains of an ICP and/or the interactive or binding partner (in cases where the binding partner is a protein), in place of one or both of the full length proteins. Any number of methods routinely practiced in the art can be used to identify and isolate the binding sites. These methods include, but are not

20 limited to, mutagenesis of the gene encoding one of the proteins and screening for disruption of binding in a co-immunoprecipitation assay. Compensatory mutations in the gene encoding the second species in the complex can then be selected. Sequence analysis of the genes encoding the respective proteins will reveal the mutations that correspond to the region of the protein involved in interactive binding. Alternatively, one protein can be

25 anchored to a solid surface using methods described above, and allowed to interact with and bind to its labeled binding partner, which has been treated with a proteolytic enzyme, such as trypsin. After washing, a relatively short, labeled peptide comprising the binding domain may remain associated with the solid material, which can be isolated and identified by amino acid sequencing. Also, once the gene coding for the intracellular binding partner is

30 obtained, short gene segments can be engineered to express peptide fragments of the protein, which can then be tested for binding activity and purified or synthesized.

For example, and not by way of limitation, an ICP product can be anchored to a solid material as described, above, by making a GST-ICP fusion protein and allowing it to bind to glutathione agarose beads. The interactive binding partner can be labeled with a radioactive isotope, such as  $^{35}\text{S}$ , and cleaved with a proteolytic enzyme such as trypsin. Cleavage products can then be added to the anchored GST-ICP fusion protein and allowed to bind. After washing away unbound peptides, labeled bound material, representing the intracellular binding partner binding domain, can be eluted, purified, and analyzed for amino acid sequence by well-known methods. Peptides so identified can be produced synthetically or fused to appropriate facilitative proteins using recombinant DNA technology.

## 5.6 MODULATORY, ANTISENSE, RIBOZYME AND TRIPLE HELIX APPROACHES

In another embodiment, the levels of ICP gene regulation can be reduced by using well-known antisense, gene "knock-out," ribozyme and/or triple helix methods. Such molecules may be designed to modulate, reduce or inhibit either unimpaired, or if appropriate, mutant sequence activity. Techniques for the production and use of such molecules are well known to those of skill in the art.

Antisense RNA and DNA molecules act to directly block the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense approaches involve the design of oligonucleotides which are complementary to an mRNA sequence. The antisense oligonucleotides will bind to the complementary mRNA sequence transcripts and prevent translation. Absolute complementarity, although preferred, is not required.

A sequence "complementary" to a portion of an RNA, as referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

In one embodiment, oligonucleotides complementary to non-coding regions of the sequence of interest could be used in an antisense approach to inhibit translation of endogenous mRNA. Antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific  
5 aspects, the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides or at least 50 nucleotides.

Regardless of the choice of target sequence, it is preferred that *in vitro* studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit mRNA expression. It is preferred that these studies utilize controls that distinguish between  
10 antisense gene inhibition and nonspecific biological effects of oligonucleotides. Additionally, it is envisioned that results obtained using the antisense oligonucleotide are compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleic acid of the oligonucleotide differs from the antisense sequence no more than is  
15 necessary to prevent specific hybridization to the target sequence.

The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other  
20 appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger, *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre, *et al.*, 1987, *Proc. Natl. Acad. Sci. U.S.A.* 84:648-652; PCT Publication No. WO88/09810, published December 15, 1988) or the blood-brain barrier (see, *e.g.*, PCT Publication No. WO89/10134, published April 25, 1988),  
25 hybridization-triggered cleavage agents (see, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents (see, *e.g.*, Zon, 1988, *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which  
30 is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine,

5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an  $\alpha$ -anomeric oligonucleotide. An  $\alpha$ -anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gautier, *et al.*, 1987, *Nucl. Acids Res.* 15:6625-6641). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue, *et al.*, 1987, *Nucl. Acids Res.* 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue, *et al.*, 1987, *FEBS Lett.* 215:327-330).

Oligonucleotides of the invention may be synthesized by standard methods known in the art, *e.g.*, by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein, *et al.* (1988, *Nucl. Acids Res.* 16:3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin, *et al.*, 1988, *Proc. Natl. Acad. Sci. U.S.A.* 85:7448-7451), *etc.*

While antisense nucleotides complementary to a coding region sequence could be used, those complementary to the transcribed, untranslated region are most preferred.

Antisense molecules should be delivered to cells that express the sequence *in vivo*. A number of methods have been developed for delivering antisense DNA or RNA to cells; *e.g.*, antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the desired cells (*e.g.*, antisense linked to peptides or  
5 antibodies which specifically bind receptors or antigens expressed on the target cell surface) can be administered systemically.

A preferred approach to achieve intracellular concentrations of the antisense sufficient to suppress translation of endogenous mRNAs utilizes a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol  
10 III or pol II promoter. The use of such a construct to transfect target cells in the patient will result in the transcription of sufficient amounts of single stranded RNAs which will form complementary base pairs with the endogenous sequence transcripts and thereby prevent translation of the mRNA sequence. For example, a vector can be introduced *e.g.*, such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can  
15 remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian,  
20 preferably human cells. Such promoters can be inducible or constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Bernoist and Chambon, 1981, *Nature* 290:304-310), the promoter contained in the 3'-long terminal repeat of Rous sarcoma virus (Yamamoto, *et al.*, 1980, *Cell* 22:787-797), the herpes thymidine kinase promoter (Wagner, *et al.*, 1981, *Proc. Natl. Acad. Sci. U.S.A.* 78:1441-1445), the regulatory  
25 sequences of the metallothionein gene (Brinster, *et al.*, 1982, *Nature* 296:39-42), etc. Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct that can be introduced directly into the tissue site. Alternatively, viral vectors can be used that selectively infect the desired tissue, in which case administration may be accomplished by another route (*e.g.*, systemically).

30 Ribozyme molecules designed to catalytically cleave target gene mRNA transcripts can also be used to prevent translation of target gene mRNA and, therefore, expression of

target gene product. (See, *e.g.*, PCT International Publication WO90/11364, published October 4, 1990; Sarver, *et al.*, 1990, *Science* 247, 1222-1225).

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. (For a review, see Rossi, 1994, *Current Biology* 4:469-471). The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by an endonucleolytic cleavage event. The composition of ribozyme molecules must include one or more sequences complementary to the target gene mRNA, and must include the well known catalytic sequence responsible for mRNA cleavage. For this sequence, see, *e.g.*, U.S. Patent No. 5,093,246, which is incorporated herein by reference in its entirety.

While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy target gene mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions which form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Myers, 1995, *Molecular Biology and Biotechnology: A Comprehensive Desk Reference*, VCH Publishers, New York, (see especially Figure 4, page 833) and in Haseloff and Gerlach, 1988, *Nature*, 334:585-591, which is incorporated herein by reference in its entirety.

Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the target gene mRNA, *i.e.*, to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in *Tetrahymena thermophila* (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug, *et al.*, 1984, *Science*, 224:574-578; Zaug and Cech, 1986, *Science*, 231:470-475; Zaug, *et al.*, 1986, *Nature*, 324:429-433; published International patent application No. WO 88/04300 by University Patents Inc.; Been and Cech, 1986, *Cell*, 47:207-216). The Cech-type ribozymes have an eight base pair active site that hybridizes to a target RNA sequence whereafter cleavage of the target RNA



takes place. The invention encompasses those Cech-type ribozymes that target eight base-pair active site sequences that are present in the target gene.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (*e.g.*, for improved stability, targeting, *etc.*) and should be delivered to cells that express the target gene *in vivo*. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous target gene messages and inhibit translation. Because ribozymes, unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Endogenous target gene expression can also be reduced by inactivating or "knocking out" the target gene or its promoter using targeted homologous recombination (*e.g.*, see Smithies, *et al.*, 1985, *Nature* 317:230-234; Thomas and Capecchi, 1987, *Cell* 51:503-512; Thompson, *et al.*, 1989, *Cell* 5:313-321; each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional target gene (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous target gene (either the coding regions or regulatory regions of the target gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells which express the target gene *in vivo*. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the target gene. Such approaches are particularly suited in the agricultural field where modifications to ES (embryonic stem) cells can be used to generate animal offspring with an inactive target gene (*e.g.*, see Thomas and Capecchi, 1987 and Thompson, 1989, *supra*). However this approach can be adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors.

Alternatively, endogenous target gene expression can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the target gene (*i.e.*, the target gene promoter and/or enhancers) to form triple helical structures which prevent transcription of the target gene in target cells in the body. (See generally, Helene, 1991, *Anticancer Drug Des.*, 6(6):569-584; Helene, *et al.*, 1992, *Ann. N.Y. Acad. Sci.*, 660:27-36; and Maher, 1992, *Bioassays* 14(12):807-815).

Nucleic acid molecules to be used in triple helix formation for the inhibition of transcription should be single stranded and composed of deoxynucleotides. The base composition of these oligonucleotides must be designed to promote triple helix formation via Hoogsteen base pairing rules, which generally require sizable stretches of either purines  
5 or pyrimidines to be present on one strand of a duplex. Nucleic acids may be pyrimidine-based, which will result in TAT and CGC<sup>+</sup> triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen which are purine-rich, for example, contain  
10 a stretch of G residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the purine residues are located on a single strand of the targeted duplex, resulting in GGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so-called "switchback" nucleic acid molecule. Switchback  
15 molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizable stretch of either purines or pyrimidines to be present on one strand of a duplex.

In instances wherein the antisense, ribozyme, and/or triple helix molecules described herein are utilized to inhibit mutant gene expression, it is possible that the technique may so  
20 efficiently reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles which the possibility may arise wherein the concentration of normal target gene product present may be lower than is necessary for a normal phenotype. In such cases, to ensure that substantially normal levels of target gene activity are maintained, therefore, nucleic acid molecules which encode and  
25 express target gene polypeptides exhibiting normal target gene activity may, be introduced into cells via gene therapy methods such as those described, below, which do not contain sequences susceptible to whatever antisense, ribozyme, or triple helix treatments are being utilized. Alternatively, in instances whereby the target gene encodes an extracellular protein, it may be preferable to co-administer normal target gene protein in order to maintain  
30 the requisite level of target gene activity.

Anti-sense RNA and DNA, ribozyme, and triple helix molecules of the invention may be prepared by any method known in the art for the synthesis of DNA and RNA molecules, as discussed above. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for  
5 example solid-phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors that incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense  
10 RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

#### 5.6.1 GENE REPLACEMENT THERAPY

An alternative means for employing the presently disclosed ICP agents includes the use of vectors to directly insert genes encoding the agents into target cells (e.g., gene  
15 therapy).

The nucleic acid sequences can be utilized for transferring recombinant nucleic acid sequences to cells and expressing said sequences in recipient cells. Such techniques can be used, for example, in marking cells or for the treatment of a condition, disorder, or disease. Such treatment can be in the form of gene replacement therapy. Specifically, one or more  
20 copies of a normal sequence or a portion of the sequence which directs the production of a sequence product exhibiting normal sequence function, may be inserted into the appropriate cells within a patient, using vectors which include, but are not limited to adenovirus, adeno-associated virus and retrovirus vectors, in addition to other particles which introduce DNA into cells, such as liposomes.

25 In another embodiment, techniques for delivery involve direct administration, *e.g.*, by stereotactic delivery of such sequences to the site of the cells in which the sequences are to be expressed.

Methods for introducing genes for expression in mammalian cells are well known in the field. Generally, for such gene therapy methods, the nucleic acid is directly administered  
30 *in vivo* into a target cell or a transgenic mouse that expresses SP-10 promoter operably linked to a reporter gene. This can be accomplished by any methods known in the art, *e.g.*,

by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, *e.g.*, by infection using a defective or attenuated retroviral or other viral vector (see U.S. Patent No. 4,980,286), by direct injection of naked DNA, by use of microparticle bombardment (*e.g.*, a gene gun; Biolistic, Dupont), by coating with  
5 lipids or cell-surface receptors or transfecting agents, by encapsulation in liposomes, microparticles, or microcapsules, by administering it in linkage to a peptide which is known to enter the nucleus, or by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see *e.g.*, Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432), which can be used to target cell types specifically expressing the receptors. In another  
10 embodiment, a nucleic acid-ligand complex can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted *in vivo* for cell specific uptake and expression, by targeting a specific receptor (see, *e.g.*, PCT Publications WO 92/06180 dated April 16, 1992; WO 92/22635 dated December 23, 1992; WO92/20316  
15 dated November 26, 1992; WO93/14188 dated July 22, 1993; WO 93/20221 dated October 14, 1993).

Additional methods which may be utilized to increase the overall level of expression of sequences of the invention include using targeted homologous recombination methods, discussed, above, to modify the expression characteristics of an endogenous sequence in a  
20 cell or microorganism by inserting a heterologous DNA regulatory element such that the inserted regulatory element is operatively linked with the endogenous sequence in question. Targeted homologous recombination can thus be used to activate transcription of an endogenous sequence which is "transcriptionally silent", *i.e.*, is not normally expressed or is normally expressed at very low levels, or to enhance the expression of an endogenous  
25 sequence which is normally expressed.

Further, the overall level of expression of sequences may be increased by the introduction of appropriate sequence-expressing cells, preferably autologous cells, into a patient at positions and in numbers which are sufficient to ameliorate the symptoms of a condition, disorder, or disease involving ICPs. Such cells may be either recombinant or  
30 non-recombinant.

Among the cells that can be administered to increase the overall level of sequence expression in a patient are normal cells which express the sequence. Alternatively, cells, preferably autologous cells, can be engineered to express the sequences, and may then be introduced into a patient in positions appropriate for the amelioration of the symptoms of a  
5 condition, disorder, or disease involving ICPs.

When the cells to be administered are non-autologous cells, they can be administered using well-known techniques that prevent a host immune response against the introduced cells from developing. For example, the cells may be introduced in an encapsulated form that, while allowing for an exchange of components with the immediate extracellular  
10 environment, does not allow the introduced cells to be recognized by the host immune system.

### **5.7 PHARMACEUTICAL FORMULATIONS AND METHODS OF ADMINISTRATION**

The compounds of this invention can be formulated and administered to inhibit a  
15 variety of disease states by any means that produces contact of the active ingredient with the agent's site of action in the body of a mammal. They can be administered by any conventional means available for use in conjunction with pharmaceuticals, either as individual therapeutic active ingredients or in a combination of therapeutic active ingredients. They can be administered alone, but are generally administered with a  
20 pharmaceutical carrier selected on the basis of the chosen route of administration and standard pharmaceutical practice.

The dosage administered will be a therapeutically effective amount of the compound sufficient to result in amelioration of symptoms of the disease and will, of course, vary depending upon known factors such as the pharmacodynamic characteristics of the particular  
25 active ingredient and its mode and route of administration; age, sex, health and weight of the recipient; nature and extent of symptoms; kind of concurrent treatment, frequency of treatment and the effect desired.

Preferably, agents that modulate ICP function shall be substantially specific. For the purposes of the present invention, the term substantially specific shall mean that a given  
30 agent is capable of being dosaged to provide the desired effect while not causing undue cellular toxicity.

One of ordinary skill will appreciate that, from a medical practitioner's or patient's perspective, virtually any alleviation or prevention of an undesirable symptom (e.g., symptoms related to disease, sensitivity to environmental factors, normal aging, and the like) would be desirable. Thus, for the purposes of this Application, the terms "treatment",

5 "therapeutic use", or "medicinal use" used herein shall refer to any and all uses of compositions comprising the claimed agents which remedy a disease state or symptoms, or otherwise prevent, hinder, retard, or reverse the progression of disease or other undesirable symptoms in any way whatsoever.

When used in the therapeutic treatment of disease, an appropriate dosage of presently  
10 described agents, or derivatives thereof, may be determined by any of several well established methodologies. For instance, animal studies are commonly used to determine the maximal tolerable dose, or MTD, of bioactive agent per kilogram weight. In general, at least one of the animal species tested is mammalian. Those skilled in the art regularly extrapolate doses for efficacy and avoiding toxicity to other species, including human.  
15 Before human studies of efficacy are undertaken, Phase I clinical studies in normal subjects help establish safe doses.

Additionally, the bioactive agents may be complexed with a variety of well established compounds or structures that, for instance, enhance the stability of the bioactive agent, or otherwise enhance its pharmacological properties (e.g., increase *in vivo* half-life,  
20 reduce toxicity, etc.).

### 5.7.1 DOSE DETERMINATIONS

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD<sub>50</sub> (the dose lethal to 50% of the population) and the ED<sub>50</sub> (the dose therapeutically  
25 effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD<sub>50</sub>/ED<sub>50</sub>. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected  
30 cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED<sub>50</sub> with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed  
5 and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC<sub>50</sub> (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such  
10 information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

Specific dosages may also be utilized for antibodies. Typically, the preferred dosage is 0.1 mg/kg to 100 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg), and if the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is usually appropriate. If  
15 the antibody is partially human or fully human, it generally will have a longer half-life within the human body than other antibodies. Accordingly, lower dosages of partially human and fully human antibodies is often possible. Additional modifications may be used to further stabilize antibodies. For example, lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (*e.g.*, into the brain). A method for lipidation  
20 of antibodies is described by Cruikshank *et al.* ((1997) *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193).

A therapeutically effective amount of protein or polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more  
25 preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight.

Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide or antibody can include a single treatment or, preferably, can include a series of treatments. In a preferred example, a subject is treated with antibody, protein, or  
30 polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week

for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5 or 6 weeks.

The present invention further encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small  
5 molecules include, but are not limited to, peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (*i.e.*, including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole,  
10 organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds.

It is understood that appropriate doses of small molecule agents depends upon a number of factors known to those of ordinary skill in the art, *e.g.*, a physician. The dose(s)  
15 of the small molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention. Exemplary doses include milligram or microgram amounts of the small molecule per  
20 kilogram of subject or sample weight (*e.g.*, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram).

### **5.7.2 FORMULATIONS AND USE**

Pharmaceutical compositions for use in accordance with the present invention may  
25 be formulated in conventional manner using one or more physiologically acceptable carriers or excipients.

Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral or rectal administration.



For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (*e.g.*, pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (*e.g.*, lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (*e.g.*, magnesium stearate, talc or silica); disintegrants (*e.g.*, potato starch or sodium starch glycolate); or wetting agents (*e.g.*, sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (*e.g.*, sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (*e.g.*, lecithin or acacia); non-aqueous vehicles (*e.g.*, almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (*e.g.*, methyl or propyl-*p*-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound.

For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of *e.g.* gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampoules or in multi-dose containers, with an added preservative.

The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use. In general, 5 water, a suitable oil, saline, aqueous dextrose (glucose), and related sugar solutions and glycols such as propylene glycol or polyethylene glycols are suitable carriers for parenteral solutions. Solutions for parenteral administration contain preferably a water soluble salt of the active ingredient, suitable stabilizing agents and, if necessary, buffer substances. Antioxidizing agents such as sodium bisulfate, sodium sulfite or ascorbic acid, either alone 10 or combined, are suitable stabilizing agents. Also used are citric acid and its salts and sodium ethylenediaminetetraacetic acid (EDTA). In addition, parenteral solutions can contain preservatives such as benzalkonium chloride, methyl- or propyl-paraben and chlorobutanol. Suitable pharmaceutical carriers are described in *Remington's Pharmaceutical Sciences*, a standard reference text in this field.

15 The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by 20 implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

Additionally, standard pharmaceutical methods can be employed to control the 25 duration of action. These are well known in the art and include control release preparations and can include appropriate macromolecules, for example polymers, polyesters, polyamino acids, polyvinyl, pyrrolidone, ethylenevinylacetate, methyl cellulose, carboxymethyl cellulose or protamine sulfate. The concentration of macromolecules as well as the methods of incorporation can be adjusted in order to control release. Additionally, the agent can be 30 incorporated into particles of polymeric materials such as polyesters, polyamino acids,

hydrogels, poly (lactic acid) or ethylenevinylacetate copolymers. In addition to being incorporated, these agents can also be used to trap the compound in microcapsules.

Another aspect of the present invention includes formulations that provide for the sustained release of ICP antagonists. Examples of such sustained release formulations include composites of biocompatible polymers, such as poly(lactic acid), poly(lactic-co-glycolic acid), methylcellulose, hyaluronic acid, collagen, and the like. The structure, selection and use of degradable polymers in drug delivery vehicles have been reviewed in several publications, including, A. Domb *et al.*, Polymers for Advanced Technologies 3:279-292 (1992). Additional guidance in selecting and using polymers in pharmaceutical formulations can be found in the text by M. Chasin and R. Langer (eds.), "Biodegradable Polymers as Drug Delivery Systems," Vol. 45 of "Drugs and the Pharmaceutical Sciences," M. Dekker, New York, 1990. Liposomes may also be used to provide for the sustained release of ICP antagonists. Details concerning how to use and make liposomal formulations of drugs of interest can be found in, among other places, U.S. Pat. No 4,944,948; U.S. Pat. No. 5,008,050; U.S. Pat. No. 4,921,706; U.S. Pat. No. 4,927,637; U.S. Pat. No. 4,452,747; U.S. Pat. No. 4,016,100; U.S. Pat. No. 4,311,712; U.S. Pat. No. 4,370,349; U.S. Pat. No. 4,372,949; U.S. Pat. No. 4,529,561; U.S. Pat. No. 5,009,956; U.S. Pat. No. 4,725,442; U.S. Pat. No. 4,737,323; U.S. Pat. No. 4,920,016. Sustained release formulations are of particular interest when it is desirable to provide a high local concentration of ICP antagonist.

Where diagnostic, therapeutic or medicinal use of the presently described agents, or derivatives thereof, is contemplated, the bioactive agents may be introduced *in vivo* by any of a number of established methods. For instance, the agent may be administered by inhalation; by subcutaneous (sub-q); intravenous (I.V.), intraperitoneal (I.P.), or intramuscular (I.M.) injection; or as a topically applied agent (transdermal patch, ointments, creams, salves, eye drops, and the like).

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

Useful pharmaceutical dosage forms, for administration of the compounds of this invention can be illustrated as follows:

Capsules: Capsules are prepared by filling standard two-piece hard gelatin capsules each with the desired amount of powdered active ingredient, 175 milligrams of lactose, 24 milligrams of talc and 6 milligrams magnesium stearate.

Soft Gelatin Capsules: A mixture of active ingredient in soybean oil is prepared and  
5 injected by means of a positive displacement pump into gelatin to form soft gelatin capsules containing the desired amount of the active ingredient. The capsules are then washed and dried.

Tablets: Tablets are prepared by conventional procedures so that the dosage unit is the desired amount of active ingredient. 0.2 milligrams of colloidal silicon dioxide, 5  
10 milligrams of magnesium stearate, 275 milligrams of microcrystalline cellulose, 11 milligrams of cornstarch and 98.8 milligrams of lactose. Appropriate coatings may be applied to increase palatability or to delay absorption.

Injectable: A parenteral composition suitable for administration by injection is prepared by stirring 1.5% by weight of active ingredients in 10% by volume propylene  
15 glycol and water. The solution is made isotonic with sodium chloride and sterilized.

Suspension: An aqueous suspension is prepared for oral administration so that each 5 millimeters contain 100 milligrams of finely divided active ingredient, 200 milligrams of sodium carboxymethyl cellulose, 5 milligrams of sodium benzoate, 1.0 grams of sorbitol solution U.S.P. and 0.025 millimeters of vanillin.

20 Gene Therapy Administration: Where appropriate, the gene therapy vectors can be formulated into preparations in solid, semisolid, liquid or gaseous forms such as tablets, capsules, powders, granules, ointments, solutions, suppositories, injections, inhalants, and aerosols, in the usual ways for their respective route of administration. Means known in the art can be utilized to prevent release and absorption of the composition until it reaches the  
25 target organ or to ensure timed-release of the composition. A pharmaceutically acceptable form should be employed which does not ineffectuate the compositions of the present invention. In pharmaceutical dosage forms, the compositions can be used alone or in appropriate association, as well as in combination, with other pharmaceutically active compounds.

30 Accordingly, the pharmaceutical composition of the present invention may be delivered via various routes and to various sites in an animal body to achieve a particular

effect (see, *e.g.*, Rosenfeld *et al.* (1991), *supra*; Rosenfeld *et al.*, Clin. Res., 3 9(2), 31 1A (1991 a); Jaffe *et al.*, *supra*; Berkner, *supra*). One skilled in the art will recognize that although more than one route can be used for administration, a particular route can provide a more immediate and more effective reaction than another route. Local or systemic delivery  
5 can be accomplished by administration comprising application or instillation of the formulation into body cavities, inhalation or insufflation of an aerosol, or by parenteral introduction, comprising intramuscular, intravenous, peritoneal, subcutaneous, intradermal, as well as topical administration.

The composition of the present invention can be provided in unit dosage form  
10 wherein each dosage unit, *e.g.*, a teaspoonful, tablet, solution, or suppository, contains a predetermined amount of the composition, alone or in appropriate combination with other active agents. The term "unit dosage form" as used herein refers to physically discrete units suitable as unitary dosages for human and animal subjects, each unit containing a predetermined quantity of the compositions of the present invention, alone or in combination  
15 with other active agents, calculated in an amount sufficient to produce the desired effect, in association with a pharmaceutically acceptable diluent, carrier, or vehicle, where appropriate. The specifications for the unit dosage forms of the present invention depend on the particular effect to be achieved and the particular pharmacodynamics associated with the pharmaceutical composition in the particular host.

20 Accordingly, the present invention also provides a method of transferring a therapeutic gene to a host, which comprises administering the vector of the present invention, preferably as part of a composition, using any of the aforementioned routes of administration or alternative routes known to those skilled in the art and appropriate for a particular application. The "effective amount" of the composition is such as to produce the  
25 desired effect in a host which can be monitored using several end-points known to those skilled in the art. Effective gene transfer of a vector to a host cell in accordance with the present invention to a host cell can be monitored in terms of a therapeutic effect (*e.g.* alleviation of some symptom associated with the particular disease being treated) or, further, by evidence of the transferred gene or expression of the gene within the host (*e.g.*, using the  
30 polymerase chain reaction in conjunction with sequencing, Northern or Southern hybridizations, or transcription assays to detect the nucleic acid in host cells, or using

immunoblot analysis, antibody-mediated detection, mRNA or protein half-life studies, or particularized assays to detect protein or polypeptide encoded by the transferred nucleic acid, or impacted in level or function due to such transfer).

These methods described herein are by no means all-inclusive, and further methods  
5 to suit the specific application will be apparent to the ordinary skilled artisan. Moreover, the effective amount of the compositions can be further approximated through analogy to compounds known to exert the desired effect.

Furthermore, the actual dose and schedule can vary depending on whether the compositions are administered in combination with other pharmaceutical compositions, or  
10 depending on interindividual differences in pharmacokinetics, drug disposition, and metabolism. Similarly, amounts can vary in *in vitro* applications depending on the particular cell line utilized (*e.g.*, based on the number of adenoviral receptors present on the cell surface, or the ability of the particular vector employed for gene transfer to replicate in that cell line). Furthermore, the amount of vector to be added per cell will likely vary with the  
15 length and stability of the therapeutic gene inserted in the vector, as well as also the nature of the sequence, and is particularly a parameter which needs to be determined empirically, and can be altered due to factors not inherent to the methods of the present invention (for instance, the cost associated with synthesis). One skilled in the art can easily make any necessary adjustments in accordance with the exigencies of the particular situation.

20 The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the invention, and functionally equivalent methods and components are within the scope of the invention. Indeed, various modifications of the invention, in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing  
25 description and accompanying drawings. Such modifications are intended to fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising a nucleotide sequence that encodes an ion channel protein having the amino acid sequence of SEQ ID NO:2 or SEQ ID  
5 NO:4.
2. The isolated nucleic acid molecule of Claim 1 comprising the nucleic acid sequence of SEQ ID NO:1 or SEQ ID NO:3.
3. An isolated nucleic acid molecule comprising a complement of the nucleic  
10 acid molecule of Claim 1.
4. An isolated nucleic acid molecule comprising at least 24 contiguous bases of the nucleic acid molecule of Claim 2, or the complement thereof.
5. An isolated nucleic acid molecule comprising a nucleotide sequence that encodes at least thirty contiguous amino acids of the amino acid sequence shown in SEQ ID  
15 NO: 2 or SEQ ID NO:4.
6. An isolated nucleic acid molecule which hybridizes to the nucleic acid molecule as in one of claims 1-5 under highly stringent conditions.
7. An isolated nucleic acid molecule which hybridizes to the nucleic acid molecule as in one of claims 1-5 under moderately stringent conditions.
- 20 8. A vector comprising the nucleic acid molecule of Claim 6.
9. An expression vector comprising the nucleic acid molecule of Claim 6 operatively associated with a regulatory nucleic acid controlling the expression of the nucleic acid in a host cell.

10. A host cell genetically engineered to express the nucleic acid molecule of Claim 6.
11. A host cell genetically engineered to express the nucleic acid molecule of Claim 6 operatively associated with a regulatory nucleic acid controlling the expression of the nucleic acid in the host cell.
12. A transgenic, non-human animal, which has been genetically engineered to contain a transgene comprising the nucleic acid molecule of Claim 6.
13. An isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4.
14. An isolated polypeptide comprising an amino acid sequence encoded by the isolated nucleic acid molecule of Claim 6.
15. An isolated polypeptide comprising an amino acid sequence encoded by the isolated nucleic acid molecule of Claim 2.
16. An antibody which binds to an isolated polypeptide comprising an amino acid sequence encoded by the isolated nucleic acid molecule of Claim 6.
17. A method for identifying a compound which modulates expression of a novel ion channel protein comprising (ICP):
- (a) contacting a test compound to a cell that expresses an ICP;
  - (b) measuring a level of the ICP expression in the cell; and
  - (c) comparing the level of the ICP expression in the cell in the presence of the test compound to a level of the ICP expression in the cell in the absence of the test compound;



wherein, if the level of the ICP expression in the cell in the presence of the test compound differs from the level of expression of the ICP in the cell in the absence of the test compound, a compound that modulates expression of the ICP is identified.

18. A method for transferring a novel ion channel protein (ICP) in a cell  
5 comprising contacting the cell with a nucleic acid comprising an ICP such that the ICP is transferred into the cell.

19. The method of Claim 17 wherein the ICP is expressed in the cell.

## SEQUENCE LISTING

<110> Turner, C. Alexander, Jr.  
 Olson, Andrew  
 Zambrowicz, Brian  
 Friedrich, Glenn  
 Sands, Arthur T.

<120> Novel Human Ion Channel Proteins

<130> 8535-0042-999

<160> 4

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 1197

<212> DNA

<213> Homo sapiens

<400> 1

atgtctcaac accgtcacca ggcgcactcg agagtcattt ctagttcacc agttgacact	60
acatcggtgg gattttgccc aacattcaag aaatttaaga ggaacgatga tgaatgtcgg	120
gcattttgtga agagagtcac aatgagccgt ttctttaaga taattatgat tagcactgtc	180
acatcgaatg cgttttttat ggccttgtgg accagttatg acataaggta ccgcttggtc	240
agacttcttg agttctcgga gatcttcttt gtgtccatct gcacatctga gttgtccatg	300
aaggtctatg tggaccccat caactactgg aagaacggct acaacctgct ggatgtgatc	360
attatcatcg ttatgttttt accctatgcc ctccgccagc tcatgggcaa acagttcact	420
tacctgtata tcgctgatgg catgcagtec ctgcgcatcc tcaagcttat cggctatagc	480
cagggcaccc ggacgctgat caccgccgtg gggcagacag tctacaccgt ggcctctgtg	540
ctcctcctgc tcttcctcct catgtacatc ttcgctatct tgggcttctg cctgtttgga	600
tctccagaca atggtgacca tgataactgg gggaaacctg ctggggcttt ttccaccctc	660
ttcagcttgg ccacggttga tggctggaca gacctgcaga agcagttgga caatcgggaa	720
tttgctttga gccgggcatt caccatcacc ttcatcttgc tcgcctcttt catcttcctc	780
aacatgttcg tgggtgtgat gatcatgcac acagaggact ccatcagaaa gtttgagcga	840
gagctgatgt tggagcagca ggagatgctc atgggagaga agcaggtgat tctgcagcgg	900
cagcaggagg agatcagcag gctgatgcac atacagaaaa atgctgactg cacaagtttc	960
agtgagctgg tggagaactt taagaagacc ttgagccaca ctgacccaat ggtcttggat	1020
gattttggca ctagcttacc cttcatcgat atctactttt ccactctgga ctaccaggac	1080
acaactgtcc acaagcttca agagctgtac tatgagatcg tgcattgtgt gagcctaata	1140
ctggaagact tgccccagga gaagccccag tccttggaaa aggtggatga gaagtag	1197

<210> 2

<211> 398

<212> PRT

<213> Homo sapiens

<400> 2

Met	Ser	Gln	His	His	Gln	Arg	His	Ser	Arg	Val	Ile	Ser	Ser	Ser
1				5				10				15		
Pro	Val	Asp	Thr	Thr	Ser	Val	Gly	Phe	Cys	Pro	Thr	Phe	Lys	Lys
			20				25					30		Phe
Lys	Arg	Asn	Asp	Asp	Glu	Cys	Arg	Ala	Phe	Val	Lys	Arg	Val	Ile
														Met

35 40 45  
 Ser Arg Phe Phe Lys Ile Ile Met Ile Ser Thr Val Thr Ser Asn Ala  
 50 55 60  
 Phe Phe Met Ala Leu Trp Thr Ser Tyr Asp Ile Arg Tyr Arg Leu Phe  
 65 70 75 80  
 Arg Leu Leu Glu Phe Ser Glu Ile Phe Phe Val Ser Ile Cys Thr Ser  
 85 90 95  
 Glu Leu Ser Met Lys Val Tyr Val Asp Pro Ile Asn Tyr Trp Lys Asn  
 100 105 110  
 Gly Tyr Asn Leu Leu Asp Val Ile Ile Ile Ile Val Met Phe Leu Pro  
 115 120 125  
 Tyr Ala Leu Arg Gln Leu Met Gly Lys Gln Phe Thr Tyr Leu Tyr Ile  
 130 135 140  
 Ala Asp Gly Met Gln Ser Leu Arg Ile Leu Lys Leu Ile Gly Tyr Ser  
 145 150 155 160  
 Gln Gly Ile Arg Thr Leu Ile Thr Ala Val Gly Gln Thr Val Tyr Thr  
 165 170 175  
 Val Ala Ser Val Leu Leu Leu Leu Phe Leu Leu Met Tyr Ile Phe Ala  
 180 185 190  
 Ile Leu Gly Phe Cys Leu Phe Gly Ser Pro Asp Asn Gly Asp His Asp  
 195 200 205  
 Asn Trp Gly Asn Leu Ala Ala Phe Phe Thr Leu Phe Ser Leu Ala  
 210 215 220  
 Thr Val Asp Gly Trp Thr Asp Leu Gln Lys Gln Leu Asp Asn Arg Glu  
 225 230 235 240  
 Phe Ala Leu Ser Arg Ala Phe Thr Ile Ile Phe Ile Leu Leu Ala Ser  
 245 250 255  
 Phe Ile Phe Leu Asn Met Phe Val Gly Val Met Ile Met His Thr Glu  
 260 265 270  
 Asp Ser Ile Arg Lys Phe Glu Arg Glu Leu Met Leu Glu Gln Gln Glu  
 275 280 285  
 Met Leu Met Gly Glu Lys Gln Val Ile Leu Gln Arg Gln Gln Glu Glu  
 290 295 300  
 Ile Ser Arg Leu Met His Ile Gln Lys Asn Ala Asp Cys Thr Ser Phe  
 305 310 315 320  
 Ser Glu Leu Val Glu Asn Phe Lys Lys Thr Leu Ser His Thr Asp Pro  
 325 330 335  
 Met Val Leu Asp Asp Phe Gly Thr Ser Leu Pro Phe Ile Asp Ile Tyr  
 340 345 350  
 Phe Ser Thr Leu Asp Tyr Gln Asp Thr Thr Val His Lys Leu Gln Glu  
 355 360 365  
 Leu Tyr Tyr Glu Ile Val His Val Leu Ser Leu Met Leu Glu Asp Leu  
 370 375 380  
 Pro Gln Glu Lys Pro Gln Ser Leu Glu Lys Val Asp Glu Lys  
 385 390 395

&lt;210&gt; 3

&lt;211&gt; 1254

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 3

atgtctcaac accgtcacca gcgcactcg agagtcattt ctagttcacc agttgacact	60
acatcggtgg gattttgccc aacattcaag aaatttaaga ggaacgatga tgaatgtcgg	120
gcatttgatga agagagtcac aatgagccgt ttctttaaga taattatgat tagcactgtc	180
acatcgaatg cgttttttat ggccttggtg accagttatg acataaggta ccgcttggtc	240

```

agacttcttg aggggtgggtg ctcagtcaca gctcttccag tgaatgaatg gatttgcatg      300
atagtcaagt tctcggagat cttctttgtg tccatctgca catctgagtt gtccatgaag      360
gtctatgtgg accccatcaa ctactggaag aacggctaca acctgctgga tgtgatcatt      420
atcatcgta tgtttttacc ctatgcctc cgccagctca tgggcaaaca gttcacttac      480
ctgtatatcg ctgatggcat gcagtcctcg cgcacctca agcttatcgg ctatagccag      540
ggcatccgga cgctgacac cgccgtgggg cagacagtct acaccgtggc ctctgtgctc      600
ctcctgctct tctcctcat gtacatcttc gctatcttgg gcttctgcct gtttgatct      660
ccagacaatg gtgaccatga taactggggg aacctggctg cggtttttt caccctcttc      720
agcttggcca cggttgatgg ctggacagac ctgcagaagc agttggacaa tcgggaattt      780
gctttgagcc gggcattcac catcatcttc atcttgctcg cctctttcat ctctctcaac      840
atgttcgtgg gtgtgatgat catgcacaca gaggactcca tcagaaagt tgagcgagag      900
ctgatgttgg agcagcagga gatgctcatg ggagagaagc aggtgattct gcagcggcag      960
caggaggaga tcagcaggct gatgcacata cagaaaaatg ctgactgcac aagtttcagt     1020
gagctggtgg agaactttaa gaagaccttg agccacactg acccaatggt cttggatgat     1080
tttggcacta gcttaccctt catcgatata tacttttcca ctctggacta ccaggacaca     1140
actgtccaca agcttcaaga gctgtactat gagatcgtgc atgtgctgag cctaattgctg     1200
gaagacttgc cccaggagaa gccccagtcc ttggaaaagg tggatgagaa gtag           1254

```

&lt;210&gt; 4

&lt;211&gt; 417

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 4

```

Met Ser Gln His Arg His Gln Arg His Ser Arg Val Ile Ser Ser Ser
 1          5          10          15
Pro Val Asp Thr Thr Ser Val Gly Phe Cys Pro Thr Phe Lys Lys Phe
          20          25          30
Lys Arg Asn Asp Asp Glu Cys Arg Ala Phe Val Lys Arg Val Ile Met
          35          40          45
Ser Arg Phe Phe Lys Ile Ile Met Ile Ser Thr Val Thr Ser Asn Ala
          50          55          60
Phe Phe Met Ala Leu Trp Thr Ser Tyr Asp Ile Arg Tyr Arg Leu Phe
65          70          75          80
Arg Leu Leu Glu Gly Gly Cys Ser Val Thr Ala Leu Pro Val Asn Glu
          85          90          95
Trp Ile Cys Met Ile Val Lys Phe Ser Glu Ile Phe Phe Val Ser Ile
          100          105          110
Cys Thr Ser Glu Leu Ser Met Lys Val Tyr Val Asp Pro Ile Asn Tyr
          115          120          125
Trp Lys Asn Gly Tyr Asn Leu Leu Asp Val Ile Ile Ile Val Met
          130          135          140
Phe Leu Pro Tyr Ala Leu Arg Gln Leu Met Gly Lys Gln Phe Thr Tyr
145          150          155          160
Leu Tyr Ile Ala Asp Gly Met Gln Ser Leu Arg Ile Leu Lys Leu Ile
          165          170          175
Gly Tyr Ser Gln Gly Ile Arg Thr Leu Ile Thr Ala Val Gly Gln Thr
          180          185          190
Val Tyr Thr Val Ala Ser Val Leu Leu Leu Leu Phe Leu Leu Met Tyr
          195          200          205
Ile Phe Ala Ile Leu Gly Phe Cys Leu Phe Gly Ser Pro Asp Asn Gly
          210          215          220
Asp His Asp Asn Trp Gly Asn Leu Ala Ala Ala Phe Phe Thr Leu Phe
225          230          235          240
Ser Leu Ala Thr Val Asp Gly Trp Thr Asp Leu Gln Lys Gln Leu Asp
          245          250          255

```

Asn Arg Glu Phe Ala Leu Ser Arg Ala Phe Thr Ile Ile Phe Ile Leu  
                   260                  265                  270  
 Leu Ala Ser Phe Ile Phe Leu Asn Met Phe Val Gly Val Met Ile Met  
                   275                  280                  285  
 His Thr Glu Asp Ser Ile Arg Lys Phe Glu Arg Glu Leu Met Leu Glu  
                   290                  295                  300  
 Gln Gln Glu Met Leu Met Gly Glu Lys Gln Val Ile Leu Gln Arg Gln  
 305                  310                  315                  320  
 Gln Glu Glu Ile Ser Arg Leu Met His Ile Gln Lys Asn Ala Asp Cys  
                   325                  330                  335  
 Thr Ser Phe Ser Glu Leu Val Glu Asn Phe Lys Lys Thr Leu Ser His  
                   340                  345                  350  
 Thr Asp Pro Met Val Leu Asp Asp Phe Gly Thr Ser Leu Pro Phe Ile  
                   355                  360                  365  
 Asp Ile Tyr Phe Ser Thr Leu Asp Tyr Gln Asp Thr Thr Val His Lys  
                   370                  375                  380  
 Leu Gln Glu Leu Tyr Tyr Glu Ile Val His Val Leu Ser Leu Met Leu  
 385                  390                  395                  400  
 Glu Asp Leu Pro Gln Glu Lys Pro Gln Ser Leu Glu Lys Val Asp Glu  
                   405                  410                  415  
 Lys